

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ :		A2	(11) International Publication Number:	WO 00/29574
C12N 15/12, C07K 14/47, C12N 1/00, A61K 38/17, C07K 16/18, C12Q 1/68			(43) International Publication Date:	
			25 May 2000 (25.05.00)	
(21) International Application Number:		PCT/US99/26234		
(22) International Filing Date:		4 November 1999 (04.11.99)		
(30) Priority Data: 09/195,292		18 November 1998 (18.11.98) US		
(63) Related by Continuation (CON) or Continuation-In-Part (CIP) to Earlier Application US		09/195,292 (CIP) Filed on 18 November 1998 (18.11.98)		
(71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).		
(72) Inventors; and		Published		
(75) Inventors/Applicants (for US only): WALKER, Michael, G. [CA/US]; Unit 80, 1050 Borregas Avenue, Sunnyvale, CA 94089 (US). VOLKMUTH, Wayne [US/US]; 783 Roble Avenue #1, Menlo Park, CA 94025 (US). KLINGLER, Tod, M. [US/US]; 28 Dover Court, San Carlos, CA 94070 (US).		Without international search report and to be republished upon receipt of that report.		
(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).				
(54) Title: INFLAMMATION-ASSOCIATED GENES				
(57) Abstract				

The invention provides novel inflammation-associated genes and polypeptides encoded by those genes. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating or preventing diseases associated with inflammation.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

WO 00/29574

PCT/US99/26234

INFLAMMATION-ASSOCIATED GENES**TECHNICAL FIELD**

The invention relates to 11 inflammation-associated genes identified by their coexpression with known inflammation genes; to their corresponding polypeptides; and to the use of these biomolecules in diagnosis, prognosis, prevention and evaluation of therapies for diseases associated with inflammation.

BACKGROUND ART

Inflammation is the body's immediate, general response to wounding or infection by a pathogen. There are many complex phenomena that occur during an inflammation response. Initiation of the complement cascade, leukocyte recruitment, and leukocyte activation are three key events. In the complement cascade, a set of serum proteins, collectively called complement, non-specifically coat foreign matter. The coating proceeds in a cascade of steps using particular subsets of factors called complement components. The coated particles are then phagocytosed by macrophages or neutrophils recruited to the inflammation site. Leukocyte recruitment of monocytes and neutrophils is mediated by cytokines, which are proteins secreted by tissue at the inflammation site. Interleukin-8 (IL-8) is the primary chemoattractant cytokine responsible for recruitment in the initial stage of inflammation. In response to IL-8, monocytes and neutrophils are activated. An immediate response to activation is the expression of L-selectin and the integrins. L-selectin is a surface molecule that facilitates leukocyte binding (with relatively low affinity) to the endothelial cells lining blood vessels in the vicinity of the inflammation site. The integrins, also cell surface molecules, have stronger binding and mediate the actual extravasation of leukocytes from the blood vessel. Upon reaching the site of inflammation, receptors to the complement factors coating foreign particles are expressed on the leukocytes leading to phagocytosis and enzymatic degradation. Many genes that participate in and regulate the inflammation response are known, but many remain to be identified. Identification of currently unknown genes will provide new diagnostic and therapeutic targets for control of the inflammation response and treatment of inflammatory disorders.

The present invention provides new compositions that are useful for control of inflammation and for the diagnosis, prognosis, treatment, prevention, and evaluation of therapies for inflammatory disorders such as rheumatoid arthritis, Crohn's disease, multiple sclerosis, asthma and allergy. We have identified 11 novel inflammation-associated genes by their coexpression with known inflammation genes.

DISCLOSURE OF THE INVENTION

In one aspect, the invention provides for a substantially purified polynucleotide comprising a gene that is coexpressed with one or more known inflammation genes in a plurality of biological samples. Preferably, known inflammation genes are selected from the group consisting of CD16, L-selectin, Src-like adapter protein, IP-30, superoxidase homoenzyme subunits p67phox, p47phox, and p40phox, alpha-1-antitrypsin, Clq-A, 5-lipoxygenase activating protein, and SRC family tyrosine kinase. Preferred embodiments include (a) a polynucleotide sequence of SEQ ID NOs: 1-11; (b) a polynucleotide sequence which encodes the polypeptide sequence 12, 13, 14, 15, 16, or 17; (c) a polynucleotide sequence having at least 70% identity to the polynucleotide sequence of (a) or (b); (d) a polynucleotide sequence which is complementary to the polynucleotide sequence of (a), (b), or (c); (e) a polynucleotide sequence comprising at least 10, preferably at least 18, sequential nucleotides of the polynucleotide sequence of (a), (b), (c), or (d); and (f) a polynucleotide which hybridizes under stringent conditions to the polynucleotide of (a), (b), (c), (d) or (e). Furthermore, the invention provides an expression vector comprising any of the above described polynucleotides and host cells comprising the expression vector. Still further, the invention provides a method for treating or preventing a disease or condition associated with the altered expression of a gene that is coexpressed with one or more known inflammation genes comprising administering to a subject in need a polynucleotide described above in an amount effective for treating or preventing said disease.

In a second aspect, the invention provides a substantially purified polypeptide comprising the gene product of a gene that is coexpressed with one or more known inflammation genes in a plurality of biological samples. The known inflammation gene may be selected from the group consisting of CD16, L-selectin, Src-like adapter protein, IP-30, superoxidase homoenzyme subunits p67phox, p47phox, and p40phox, alpha-1-antitrypsin, Clq-A, 5-lipoxygenase activating protein, and SRC family tyrosine kinase. Preferred embodiments are (a) the polypeptide sequence of SEQ ID NOs: 12, 13, 14, 15, 16, or 17; (b) a polypeptide sequence having at least 85% identity to the polypeptide sequence of (a); and (c) a polypeptide sequence comprising at least 6 sequential amino acids of the polypeptide sequence of (a) or (b). Additionally, the invention provides antibodies that bind specifically to any of the above described polypeptides and a method for treating or preventing a disease or condition associated with the altered expression of a gene that is coexpressed with one or more known inflammation genes comprising administering to a subject in need such an antibody in an amount effective for treating or preventing said disease.

In another aspect, the invention provides a pharmaceutical composition comprising the polynucleotide or the polypeptide in conjunction with a suitable pharmaceutical carrier and a

WO 00/29574

PCT/US99/26234

method for treating or preventing a disease or condition associated with the altered expression of a gene that is coexpressed with one or more known inflammation genes comprising administering to a subject in need such a composition in an amount effective for treating or preventing said disease.

5 In a further aspect, the invention provides a ribozyme that cleaves a polynucleotide of the invention and a method for treating or preventing a disease or condition associated with the increased expression of a gene that is coexpressed with one or more known inflammation genes. The method comprises administering to a subject in need the ribozyme in an amount effective for treating or preventing the disease.

10 In yet a further aspect, the invention provides a method for diagnosing a disease or condition associated with the altered expression of a gene that is coexpressed with one or more known inflammation genes, wherein each known inflammation gene is selected from the group consisting of CD16, L-selectin, Src-like adapter protein, IP-30, superoxidase homoenzyme subunits p67phox, p47phox, and p40phox, alpha-1-antitrypsin, Clq-A, 5-lipoxygenase activating protein, and SRC family tyrosine kinase. The method comprises the steps of (a) providing the
15 sample comprising one of more of said coexpressed genes; (b) hybridizing the polynucleotide to said coexpressed genes under conditions effective to form one or more hybridization complexes; (c) detecting the hybridization complexes; and (d) comparing the levels of the hybridization complexes with the level of hybridization complexes in a nondiseased sample, wherein altered levels of one or more of the hybridization complexes in a diseased sample compared with the level
20 of hybridization complexes in a non-diseased sample correlates with the presence of the disease or condition.

Additionally, the invention provides antibodies that bind specifically to any of the above described polypeptides and a method for treating or preventing a disease or condition associated with the inflammation response.

25

BRIEF DESCRIPTION OF THE SEQUENCE LISTING

The Sequence Listing provides exemplary inflammation-associated sequences including polynucleotide sequences, SEQ ID NOs:1-11, and polypeptide sequences, SEQ ID NOs:12-17. Each sequence is identified by a sequence identification number (SEQ ID NO) and by the Incyte Clone number from which the sequence was first identified.

30

DESCRIPTION OF THE INVENTION

It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include the plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those

WO 00/29574

PCT/US99/26234

skilled in the art, and so forth.

DEFINITIONS

"NSEQ" refers generally to a polynucleotide sequence of the present invention including SEQ ID NOs:1-11. "PSEQ" refers generally to a polypeptide sequence of the present invention including SEQ ID NOs:12-17.

A "variant" refers to either a polynucleotide or a polypeptide whose sequence diverges from SEQ ID NOs:1-11 or SEQ ID NOs:12-17, respectively. Polynucleotide sequence divergence may result from mutational changes such as deletions, additions, and substitutions of one or more nucleotides; it may also occur because of differences in codon usage. Each of these types of changes may occur alone, or in combination, one or more times in a given sequence. Polypeptide variants include sequences that possess at least one structural or functional characteristic of SEQ ID NOs:12-17.

A "fragment" can refer to a nucleic acid sequence that is preferably at least 20 nucleic acids in length, more preferably 40 nucleic acids, and most preferably 60 nucleic acids in length, and encompasses, for example, fragments consisting of nucleic acids 1-50 of SEQ ID NOs:1-11. A "fragment" can also refer to polypeptide sequences which are preferably at least 5 to about 15 amino acids in length, most preferably at least 10 amino acids, and which retain some biological or immunological activity of, for example, a sequence selected from SEQ ID NOs:12-17.

"Gene" or "gene sequence" refers to the partial or complete coding sequence of a gene.

The term also refers to 5' or 3' untranslated regions. The gene may be in a sense or antisense (complementary) orientation.

"Known inflammation gene" refers to a gene sequence which has been previously identified as useful in the diagnosis, treatment, prognosis, or prevention of diseases associated with inflammation. Typically, this means that the known gene is expressed at higher levels in tissue abundant in known inflammation transcripts when compared with other tissue.

"Inflammation-associated gene" refers to a gene sequence whose expression pattern is similar to that of the known inflammation genes and which are useful in the diagnosis, treatment, prognosis, or prevention of diseases associated with inflammation, particularly disorders such as rheumatoid arthritis, Crohn's disease, multiple sclerosis, asthma and allergy.

"Substantially purified" refers to a nucleic acid or an amino acid sequence that is removed from its natural environment and is isolated or separated, and is at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which it is naturally present.

WO 00/29574

PCT/US99/26234

THE INVENTION

The present invention encompasses a method for identifying biomolecules that are associated with a specific disease, regulatory pathway, subcellular compartment, cell type, tissue type, or species. In particular, the method identifies gene sequences useful in diagnosis, 5 prognosis, treatment, prevention, and evaluation of therapies for diseases associated with inflammation, particularly disorders such as rheumatoid arthritis, Crohn's disease, multiple sclerosis, asthma and allergy.

The method entails first identifying polynucleotides that are expressed in a plurality cDNA libraries. The identified polynucleotides include genes of known function, genes known to be 10 specifically expressed in a specific disease process, subcellular compartment, cell type, tissue type, or species. Additionally, the polynucleotides include genes of unknown function. The expression patterns of the known genes are then compared with those of the genes of unknown function to determine whether a specified coexpression probability threshold is met. Through this 15 comparison, a subset of the polynucleotides having a high coexpression probability with the known genes can be identified. The high coexpression probability correlates with a particular coexpression probability threshold which is less than 0.001, and more preferably less than 0.00001.

The polynucleotides originate from cDNA libraries derived from a variety of sources including, but not limited to, eukaryotes such as human, mouse, rat, dog, monkey, plant, and yeast 20 and prokaryotes such as bacteria and viruses. These polynucleotides can also be selected from a variety of sequence types including, but not limited to, expressed sequence tags (ESTs), assembled polynucleotide sequences, full length gene coding regions, introns, regulatory sequences, 5' untranslated regions, and 3' untranslated regions. To have statistically significant analytical results, the polynucleotides need to be expressed in at least three cDNA libraries.

25 The cDNA libraries used in the coexpression analysis of the present invention can be obtained from blood vessels, heart, blood cells, cultured cells, connective tissue, epithelium, islets of Langerhans, neurons, phagocytes, biliary tract, esophagus, gastrointestinal system, liver, pancreas, fetus, placenta, chromaffin system, endocrine glands, ovary, uterus, penis, prostate, seminal vesicles, testis, bone marrow, immune system, cartilage, muscles, skeleton, central 30 nervous system, ganglia, neuroglia, neurosecretory system, peripheral nervous system, bronchus, larynx, lung, nose, pleurus, ear, eye, mouth, pharynx, exocrine glands, bladder, kidney, ureter, and the like. The number of cDNA libraries selected can range from as few as 3 to greater than 10,000. Preferably, the number of the cDNA libraries is greater than 500.

In a preferred embodiment, gene sequences are assembled to reflect related sequences,

WO 00/29574

PCT/US99/26234

such as assembled sequence fragments derived from a single transcript. Assembly of the polynucleotide sequences can be performed using sequences of various types including, but not limited to, ESTs, extensions, or shotgun sequences. In a most preferred embodiment, the polynucleotide sequences are derived from human sequences that have been assembled using the 5 algorithm disclosed in "Database and System for Storing, Comparing and Displaying Related Biomolecular Sequence Information", Lincoln et al., Serial No:60/079,469, filed March 26, 1998, incorporated herein by reference.

Experimentally, differential expression of the polynucleotides can be evaluated by methods including, but not limited to, differential display by spatial immobilization or by gel 10 electrophoresis, genome mismatch scanning, representational difference analysis, and transcript imaging. Additionally, differential expression can be assessed by microarray technology. These methods may be used alone or in combination.

Known inflammation genes can be selected based on the use of the genes as diagnostic or prognostic markers or as therapeutic targets for diseases associated with inflammation. 15 Preferably, the known inflammation genes include CD16, L-selectin, Src-like adapter protein (SLAP), IP-30, superoxidase homoenzyme subunits (p67phox, p47phox, and p40phox), alpha-1-antitrypsin (AAT), Clq-A, 5-lipoxygenase activating protein (FLAP), and SRC family tyrosine kinase (HCK), and the like.

The procedure for identifying novel genes that exhibit a statistically significant 20 coexpression pattern with known inflammation genes is as follows. First, the presence or absence of a gene sequence in a cDNA library is defined: a gene is present in a cDNA library when at least one cDNA fragment corresponding to that gene is detected in a cDNA sample taken from the library, and a gene is absent from a library when no corresponding cDNA fragment is detected in the sample.

Second, the significance of gene coexpression is evaluated using a probability method to 25 measure a due-to-chance probability of the coexpression. The probability method can be the Fisher exact test, the chi-squared test, or the kappa test. These tests and examples of their applications are well known in the art and can be found in standard statistics texts (Agresti (1990) Categorical Data Analysis, John Wiley & Sons, New York NY; Rice (1988) Mathematical Statistics and Data Analysis, Duxbury Press, Pacific Grove CA). A Bonferroni correction (Rice, supra, page 384) can also be applied in combination with one of the probability methods for 30 correcting statistical results of one gene versus multiple other genes. In a preferred embodiment, the due-to-chance probability is measured by a Fisher exact test, and the threshold of the due-to-chance probability is set to less than 0.001, more preferably less than 0.00001.

WO 00/29574

PCT/US99/26234

To determine whether two genes, A and B, have similar coexpression patterns, occurrence data vectors can be generated as illustrated in Table 1, wherein a gene's presence is indicated by a one and its absence by a zero. A zero indicates that the gene did not occur in the library, and a one indicates that it occurred at least once.

5

Table 1. Occurrence data for genes A and B

	Library 1	Library 2	Library 3	...	Library N
gene A	1	1	0	...	0
gene B	1	0	1	...	0

10 For a given pair of genes, the occurrence data in Table 1 can be summarized in a 2 x 2 contingency table.

Table 2. Contingency table for co-occurrences of genes A and B

	Gene A present	Gene A absent	Total
Gene B present	8	2	10
Gene B absent	2	18	20
Total	10	20	30

Table 2 presents co-occurrence data for gene A and gene B in a total of 30 libraries. Both gene A and gene B occur 10 times in the libraries. Table 2 summarizes and presents 1) the

20 number of times gene A and B are both present in a library, 2) the number of times gene A and B are both absent in a library, 3) the number of times gene A is present while gene B is absent, and 4) the number of times gene B is present while gene A is absent. The upper left entry is the number of times the two genes co-occur in a library, and the middle right entry is the number of times neither gene occurs in a library. The off diagonal entries are the number of times one gene

25 occurs while the other does not. Both A and B are present eight times and absent 18 times, gene A is present while gene B is absent two times, and gene B is present while gene A is absent two times. The probability ("p-value") that the above association occurs due to chance as calculated using a Fisher exact test is 0.0003. Associations are generally considered significant if a p-value is less than 0.01 (Agresti, *supra*; Rice, *supra*).

30 This method of estimating the probability for coexpression of two genes makes several assumptions. The method assumes that the libraries are independent and are identically sampled. However, in practical situations, the selected cDNA libraries are not entirely independent because more than one library may be obtained from a single patient or tissue, and they are not entirely

WO 00/29574

PCT/US99/26234

identically sampled because different numbers of cDNA's may be sequenced from each library (typically ranging from 5,000 to 10,000 cDNA's per library). In addition, because a Fisher exact coexpression probability is calculated for each gene versus 41,419 other genes, a Bonferroni correction for multiple statistical tests is necessary.

5 Using the method of the present invention, we have identified 11 novel genes that exhibit strong association, or coexpression, with known genes that are inflammation-specific. These known inflammation genes include CD16, L-selectin, Src-like adapter protein (SLAP), IP-30, superoxidase homoenzyme subunits (p67phox, p47phox, and p40phox), alpha-1-antitrypsin (AAT), Clq-A, 5-lipoxygenase activating protein (FLAP), and SRC family tyrosine kinase (HCK).

10 The results presented in Table 5 show that the expression of the 11 novel genes have direct or indirect association with the expression of known inflammation genes. Therefore, the novel genes can potentially be used in diagnosis, treatment, prognosis, or prevention of diseases associated with inflammation, or in the evaluation of therapies for diseases associated with inflammation. Further, the gene products of the 11 novel genes are potential therapeutic proteins

15 and targets of therapeutics against diseases associated with inflammation.

Therefore, in one embodiment, the present invention encompasses a polynucleotide sequence comprising the sequence of SEQ ID Nos:1-11. These 11 polynucleotides are shown by the method of the present invention to have strong coexpression association with known inflammation genes and with each other. The invention also encompasses a variant of the 20 polynucleotide sequence, its complement, or 18 consecutive nucleotides of a sequence provided in the above described sequences. Variant polynucleotide sequences typically have at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to NSEQ. One preferred method for identifying variants entails using NSEQ and/or PSEQ sequences to search against the GenBank primate (pri), rodent (rod), and 25 mammalian (mam), vertebrate (vrtp), and eukaryote (eukp) databases, SwissProt, BLOCKS (Bairoch et al. (1997) Nucleic Acids Res 25:217-221), PFAM, and other databases that contain previously identified and annotated motifs, sequences, and gene functions. Methods that search for primary sequence patterns with secondary structure gap penalties (Smith (1992) Prot Eng 5:35-51) as well as algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul (1993) J 30 Mol Evol 36:290-300; and Altschul et al. (1990) J Mol Biol 215:403-410), BLOCKS (Henikoff and Henikoff (1991) Nucleic Acids Res 19:6565-6572), Hidden Markov Models (HMM; Eddy (1996) Cur Opin Str Biol 6:361-365; and Sonnhammer et al. (1997) Proteins 28:405-420), and the like, can be used to manipulate and analyze nucleotide and amino acid sequences. These databases, algorithms and other methods are well known in the art and are described in Ausubel et

WO 00/29574

PCT/US99/26234

al. (1997; Short Protocols in Molecular Biology, John Wiley & Sons, New York NY) and in Meyers (1995; Molecular Biology and Biotechnology, Wiley VCH, New York NY).

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to SEQ ID NOs:1-11, and fragments thereof under stringent conditions. Stringent conditions can be defined by salt concentration, temperature, and other chemicals and conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, or raising the hybridization temperature.

For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent or solvent, and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Additional variations on these conditions will be readily apparent to those skilled in the art (Wahl and Berger (1987) Methods Enzymol 152:399-407; Kimmel (1987) Methods Enzymol 152:507-511; Ausubel, supra; and Sambrook et al. (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Press, Plainview NY).

NSEQ or the polynucleotide sequences encoding PSEQ can be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. (See, e.g., Dieffenbach and Dveksler (1995) PCR Primer. a Laboratory Manual, Cold Spring Harbor Press, Plainview NY; Sarkar (1993) PCR Methods Applic 2:318-322; Triglia et al. (1988) Nucleic Acids Res 16:8186; Lagerstrom et al. (1991) PCR Methods Applic 1:111-119; and Parker et al. (1991) Nucleic Acids Res 19:3055-306). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto, CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 18 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

In another aspect of the invention, NSEQ or the polynucleotide sequences encoding PSEQ can be cloned in recombinant DNA molecules that direct expression of PSEQ or the polypeptides encoded by NSEQ, or structural or functional fragments thereof, in appropriate host cells. Due to

WO 00/29574

PCT/US99/26234

the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express the polypeptides of PSEQ or the polypeptides encoded by NSEQ. The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter the

5 nucleotide sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter
10 glycosylation patterns, change codon preference, produce splice variants, and so forth.

In order to express a biologically active polypeptide encoded by NSEQ or the polynucleotide sequences encoding PSEQ, or derivatives thereof, may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These
15 elements include regulatory sequences such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and NSEQ or polynucleotide sequences encoding PSEQ. Methods which are well known to those skilled in the art may be used to construct expression vectors containing NSEQ or polynucleotide sequences encoding PSEQ and appropriate transcriptional and translational control elements. These methods include in vitro recombinant
20 DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, supra and Ausubel, supra.)

A variety of expression vector/host cell systems may be utilized to contain and express NSEQ or polynucleotide sequences encoding PSEQ. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid
25 DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (baculovirus); plant cell systems transformed with viral expression vectors, cauliflower mosaic virus (CaMV) or tobacco mosaic virus (TMV), or with bacterial expression vectors (Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed. For long term production of recombinant proteins in
30 mammalian systems, stable expression of a polypeptide encoded by NSEQ in cell lines is preferred. For example, NSEQ or sequences encoding PSEQ can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector.

In general, host cells that contain NSEQ and that express PSEQ may be identified by a

WO 00/29574

PCT/US99/26234

variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences. Immunological methods for 5 detecting and measuring the expression of PSEQ using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS).

Host cells transformed with NSEQ or polynucleotide sequences encoding PSEQ may be 10 cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of NSEQ or polynucleotides encoding PSEQ may be designed to contain signal sequences which direct secretion of PSEQ through a prokaryotic or 15 eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" 20 form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC, Manassas, MD) and may be chosen to ensure the correct modification and processing of the foreign protein.

25 In another embodiment of the invention, natural, modified, or recombinant NSEQ or nucleic acid sequences encoding PSEQ are ligated to a heterologous sequence resulting in translation of a fusion protein containing heterologous protein moieties in any of the aforementioned host systems. Such heterologous protein moieties facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited 30 to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, c-myc, hemagglutinin (HA) and monoclonal antibody epitopes.

In another embodiment, NSEQ or sequences encoding PSEQ are synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers et al. (1980) Nucleic

WO 00/29574

PCT/US99/26234

Acids Symp Ser (7)215-223; Horn et al. (1980) Nucleic Acids Symp Ser (7)225-232; and Ausubel, *supra*). Alternatively, PSEQ or a polypeptide sequence encoded by NSEQ itself, or a fragment thereof, may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge et al. (1995) Science 269:202-204). Automated synthesis may be achieved using the ABI 431A Peptide synthesizer (PE Biosystems, Foster City CA). Additionally, PSEQ or the amino acid sequence encoded by NSEQ, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a polypeptide variant.

In another embodiment, the invention entails a substantially purified polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NOs:12-17 or fragments thereof.

DIAGNOSTICS and THERAPEUTICS

The sequences of the these genes can be used in diagnosis, prognosis, treatment, prevention, and evaluation of therapies for diseases associated with inflammation, particularly disorders such as rheumatoid arthritis, Crohn's disease, multiple sclerosis, asthma, and allergy.

In one preferred embodiment, the polynucleotide sequences of NSEQ or the polynucleotides encoding PSEQ are used for diagnostic purposes to determine the absence, presence, and excess expression of PSEQ. The polynucleotides may be at least 18 nucleotides long, complementary RNA and DNA molecules, branched nucleic acids, and peptide nucleic acids (PNAs). Alternatively, the polynucleotides are used to detect and quantitate gene expression in samples in which expression of NSEQ or the polypeptides encoded by NSEQ are correlated with disease. Additionally, NSEQ or the polynucleotides encoding PSEQ can be used to detect genetic polymorphisms associated with a disease. These polymorphisms may be detected at the transcript cDNA or genomic level.

The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding PSEQ, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at least 70% sequence identity to any of the NSEQ or PSEQ-encoding sequences.

Means for producing specific hybridization probes for DNAs encoding PSEQ include the cloning of NSEQ or polynucleotide sequences encoding PSEQ into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to

WO 00/29574

PCT/US99/26234

synthesize RNA probes in vitro by means of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, by fluorescent labels, and the like. The polynucleotide sequences encoding PSEQ may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; and in microarrays utilizing fluids or tissues from patients to detect altered PSEQ expression. Such qualitative or quantitative methods are well known in the art.

NSEQ or the nucleotide sequences encoding PSEQ can be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed, and the signal is quantitated and compared with a standard value, typically, derived from a non-diseased sample. If the amount of signal in the patient sample is altered in comparison to the standard value then the presence of altered levels of nucleotide sequences of NSEQ or those encoding PSEQ in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

Once the presence of a disease is established and a treatment protocol is initiated, hybridization or amplification assays can be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in a healthy subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

The polynucleotides may be used for the diagnosis of a variety of diseases associated with inflammation, particularly for diseases including, but not limited to, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus,

WO 00/29574

PCT/US99/26234

systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma.

5 The polynucleotides may also be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify splice variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disease, to diagnose a disease, and to develop and monitor the activities of therapeutic agents.

10 In yet another alternative, polynucleotides may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Fluorescent *in situ* hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich et al. (1995) In: Meyers, *supra*). Microarrays may be used to detect genetic diversity at the genome level.

15 In another embodiment, antibodies which specifically bind PSEQ may be used for the diagnosis of diseases characterized by the over-or-underexpression of PSEQ or polypeptides encoded by NSEQ. A variety of protocols for measuring PSEQ, or the polypeptides encoded by NSEQ, including ELISAs, RIAs, and FACS are well known in the art and provide a basis for diagnosing altered or abnormal levels of the expression of PSEQ or the polypeptides encoded by NSEQ. Standard values for PSEQ expression are established by combining body fluids or cell extracts taken from healthy subjects, preferably human, with antibody to PSEQ or a polypeptide encoded by NSEQ under conditions suitable for complex formation. The amount of complex formation may be quantitated by various methods, preferably by photometric means. Quantities of PSEQ or the polypeptides encoded by NSEQ expressed in disease samples from, for example, biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing or monitoring disease. Alternatively, one may use competitive drug screening assays in which neutralizing antibodies capable of binding PSEQ or the polypeptides encoded by NSEQ specifically compete with a test compound for binding the polypeptides. Antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PSEQ or the polypeptides encoded by NSEQ.

20 30 In another aspect, the polynucleotides and polypeptides of the present invention can be employed for treatment or the monitoring of therapeutic treatments for cancers. The polynucleotides of NSEQ or those encoding PSEQ, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotides of NSEQ or those encoding PSEQ may be used in situations in which it would be desirable to block the

WO 00/29574

PCT/US99/26234

transcription or translation of the mRNA.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art

5 can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding PSEQ. (See, e.g., Sambrook, supra; and Ausubel, supra.)

Genes having polynucleotide sequences of NSEQ or those encoding PSEQ can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or a fragment thereof, encoding PSEQ. Such constructs may be used to introduce

10 untranslatable sense or antisense sequences into a cell. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent
15 therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee et al. (1994) In: Huber and Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.)

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the cleavage of mRNA and decrease the levels of particular mRNAs, such as those comprising the polynucleotide
20 sequences of the invention. (See, e.g. Rossi, (1994) Current Biology 4:469-471.) Ribozymes may cleave mRNA at specific cleavage sites. Alternatively, ribozymes may cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The construction and production of ribozymes is well known in the art and is described in Meyers
(supra).

25 RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiester linkages within the backbone of the molecule. Alternatively, nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of
30 adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases may be included.

Alternatively, the polynucleotides of the invention may be integrated into a genome by somatic or germ cell gene therapy. Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors

WO 00/29574

PCT/US99/26234

may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman et al. (1997) *Nature Biotechnol* 15:462-466.)

5 Additionally, endogenous polynucleotide expression may be inactivated using homologous recombination methods which insert an inactive gene sequence at the target sequence location. (See, e.g. Thomas and Capecchi (1987) *Cell* 51:503-512.)

Further, an antagonist or antibody of a polypeptide of PSEQ or encoded by NSEQ may be administered to a subject to treat or prevent a cancer associated with increased expression or
10 activity of PSEQ. An antibody which specifically binds the polypeptide may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express the the polypeptide.

Antibodies to PSEQ may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single
15 chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies, those which inhibit dimer formation, are especially preferred for therapeutic use. Monoclonal antibodies to PSEQ may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique,
20 and the EBV-hybridoma technique. In addition, techniques developed for the production of chimeric antibodies can be used. (See, e.g., Meyers, *supra*.) Alternatively, techniques described for the production of single chain antibodies may be employed. Antibody fragments which contain specific binding sites for PSEQ or the polypeptide sequences encoded by NSEQ may also be generated.

25 Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art.

Yet further, an agonist of PSEQ or that encoded by NSEQ may be administered to a
30 subject to treat or prevent a cancer associated with decreased expression or activity of the polypeptide.

An additional aspect of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of

WO 00/29574

PCT/US99/26234

polypeptides of PSEQ or those encoded by NSEQ, antibodies to the polypeptides, and mimetics, agonists, antagonists, or inhibitors of the polypeptides. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example, polypeptides of PSEQ or those encoded by NSEQ, or fragments thereof, antibodies of the polypeptides, and agonists, antagonists or inhibitors of the polypeptides, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics.

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

EXAMPLES

It is understood that this invention is not limited to the particular methodology, protocols, and reagents described, as these may vary. It is also understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to limit the scope

WO 00/29574

PCT/US99/26234

of the present invention which will be limited only by the appended claims. The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention.

I cDNA Library Construction

5 The cDNA library, OVARTUT05, was selected as an example to demonstrate the construction of the cDNA libraries from which the sequences used to identify genes associated with inflammation were derived. The OVARTUT05 library was constructed from tumorous ovary tissue obtained from a 62 year-old Caucasian female. Pathology indicated a grade 4 endometrioid carcinoma with extensive squamous differentiation forming a solid mass in the right ovary. The
10 cervix showed mild chronic cervicitis, and the posterior uterine serosa showed focal endometriosis. Prior pathology indicated weakly proliferative endometrium with excessive stromal breakdown in the uterus and a mild chronic cervicitis with prominent nabothian cyst in the cervix.

The frozen tissue was homogenized and lysed using a POLYTRON homogenizer
15 (PT-3000; Brinkmann Instruments, Westbury NY) in guanidinium isothiocyanate solution. The lysate was centrifuged over a 5.7 M CsCl cushion using an SW28 rotor in a L8-70M ultracentrifuge (Beckman Coulter, Fullerton CA) for 18 hours at 25,000 rpm at ambient temperature. The RNA was extracted with acid phenol, pH 4.7, precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol, resuspended in RNase-free water, and treated with DNase at
20 37°C. The RNA extraction was repeated with acid phenol, pH 4.7 and precipitated with sodium acetate and ethanol as before. The mRNA was then isolated using the OLIGOTEX kit (Qiagen, Chatsworth CA) and used to construct the cDNA library.

The mRNA was handled according to the recommended protocols in the SUPERSCRIPT Plasmid system (Life Technologies, Rockville MD). The cDNAs were fractionated on a
25 SEPHAROSE CL4B column (Amersham Pharmacia Biotech, Piscataway NJ), and those cDNAs exceeding 400 bp were ligated into pINCY 1 plasmid (Incyte Pharmaceuticals, Palo Alto CA). The plasmid was subsequently transformed into DH5α competent cells (Life Technologies).

II Isolation and Sequencing of cDNA Clones

Plasmid DNA was released from the cells and purified using the REAL Prep 96 Plasmid
30 kit (Qiagen). This kit enabled the simultaneous purification of 96 samples in a 96-well block using multi-channel reagent dispensers. The recommended protocol was employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile Terrific Broth (Life Technologies) with carbenicillin at 25 mg/L and glycerol at 0.4%; 2) after inoculation, the cultures were incubated for 19 hours and at the end of incubation, the cells were lysed with 0.3 ml of lysis

WO 00/29574

PCT/US99/26234

buffer; and 3) following isopropanol precipitation, the plasmid DNA pellet was resuspended in 0.1 ml of distilled water. After the last step in the protocol, samples were transferred to a 96-well block for storage at 4° C.

5 The cDNAs were prepared using a MICROLAB 2200 (Hamilton, Reno NV) in combination with DNA ENGINE thermal cyclers (PTC200; MJ Research, Watertown MA) and sequenced by the method of Sanger et al. (1975, J Mol Biol 94:441f) using ABI PRISM 377 DNA Sequencing systems (PE Biosystems).

III Selection, Assembly, and Characterization of Sequences

10 The sequences used for coexpression analysis were assembled from EST sequences, 5' and 3' longread sequences, and full length coding sequences. Selected assembled sequences were expressed in at least three cDNA libraries.

15 The assembly process is described as follows. EST sequence chromatograms were processed and verified. Quality scores were obtained using PHRED (Ewing et al. (1998) Genome Res 8:175-185; Ewing and Green (1998) Genome Res 8:186-194). Then the edited sequences were loaded into a relational database management system (RDBMS). The EST sequences were clustered into an initial set of bins using BLAST with a product score of 50. All clusters of two or more sequences were created as bins. The overlapping sequences represented in a bin correspond to the sequence of a transcribed gene.

20 Assembly of the component sequences within each bin was performed using a modification of PHRAP, a publicly available program for assembling DNA fragments (Green, University of Washington, Seattle WA). Bins that showed 82% identity from a local pair-wise alignment between any of the consensus sequences were merged.

25 Bins were annotated by screening the consensus sequence in each bin against public databases, such as GBpri and GenPept from NCBI. The annotation process involved a FASTn screen against the GBpri database in GenBank. Those hits with a percent identity of greater than or equal to 70% and an alignment length of greater than or equal to 100 base pairs were recorded as homolog hits. The residual unannotated sequences were screened by FASTx against GenPept. Those hits with an E value of less than or equal to 10⁻⁸ are recorded as homolog hits.

30 Sequences were then reclustered using BLASTn and Cross-Match, a program for rapid protein and nucleic acid sequence comparison and database search (Green, *supra*), sequentially. Any BLAST alignment between a sequence and a consensus sequence with a score greater than 150 was realigned using cross-match. The sequence was added to the bin whose consensus sequence gave the highest Smith-Waterman score amongst local alignments with at least 82% identity. Non-matching sequences created new bins. The assembly and consensus generation

WO 00/29574

PCT/US99/26234

processes were performed for the new bins.

IV Coexpression Analyses of Known Inflammation Genes

Known inflammation genes were selected to identify novel genes that are closely associated with inflammation. These known genes were CD16, L-selectin, Src-like adapter protein (SLAP), IP-30, superoxidase holoenzyme subunits (p67phox, p47phox, and p40phox), alpha-1-antitrypsin (AAT), Clq-A, 5-lipoxygenase activating protein (FLAP), and SRC family tyrosine kinase (HCK).

The known inflammation genes that we examined in this analysis, and brief descriptions of their functions, are listed in Table 4.

Table 4. Known Inflammation Genes.

Gene	Description & references
CD16	receptor for IgG a.k.a. Fc gamma RIII Phagocytosis of complement-generated immune complexes occurs through CD16 (Tamm et al. (1996) J Biol Chem 271:3659-66; Marsh et al. (1998) J Immunol 160:3942-8)
L-selectin	Leukocyte adhesion molecule Binds carbohydrate ligand on endothelial cell glycoprotein Adhesion is required for extravasation near inflammation site; auxiliary function in neutrophil activation during inflammation (Frohlich et al. (1998) Blood 91:2558-64; Girard and Amalric (1998) Adv Exp Med Biol 435:55-62)
SLAP	Src-like adapter protein Associated with Eck RPTK transduction pathway; Eck RPTK autocrine loop implicated in inflammation (Pandey, et al. (1995) J Biol Chem 270:19201-4; Pandey et al. (1995) Science 268:567-9)
IP-30	Functions in MHC Class II processing of peptides; implicated in inflammation, α -interferon inducible (Luster et al. (1988) J Biol Chem 263:12036-43; Arunachalam et al. (1998) J Immunol 160:5797-806; Schuelke et al. (1998) Biochem Biophys Res Commun 245:599-606)
p67phox,	superoxidase holoenzyme subunits
p47phox,	Macrophages utilize reactive superoxide in degradation of
p40phox	phagocytosed matter, induced by complement (Ratnam and Mookerjea (1998) Immunology 94:560-568)
AAT	alpha-1-antitrypsin (AAT) inhibits trypsin, a protease

WO 00/29574

PCT/US99/26234

Differentially expressed in inflammation, several alleles linked
 to chronic inflammatory disorders (Breit and Penny (1980) Aust NZJ
 Med 10:449-53; Takeuchi, Kobayashi et al. (1984) Int J Tissue
 React 6:1-8)

5 Clq-A First complement component, subcomponent q, subunit A
 (Alberts et al. (1994) Molecular Biology of the Cell, Garland
 Publishing, New York NY, p. 1214)

FLAP S-lipoxygenase activating protein
 Lipoxygenase enzyme catalyzes formation of leukotrienes which are
 10 potent inflammatory mediators, FLAP is an anti-inflammatory
 therapeutic target (Byrum et al. (1997) J Exp Med 185:1065-75;
 Muller-Peddinghaus (1997) J Physiol Pharmacol 48:529-36)

HCK Src-family tyrosine kinase specific to hematopoietic cells
 Functions in integrin signaling, mouse knockouts have impaired
 15 inflammation response (Lowell and Berton (1998) Proc Natl Acad Sci
 95:7580-4)

The coexpression of the 11 known genes with each other is shown in Table 5. The entries in Table 5 are the negative log of the p-value (- log p) for the coexpression of the two genes. As shown, the method successfully identified the strong association of the known genes among
 20 themselves, indicating that the coexpression analysis method of the present invention was effective in identifying genes that are closely associated with inflammation.

We have also identified 11 novel genes that show strong association with known inflammation genes from a total of 41,419 assembled gene sequences. The degree of association was measured by probability values and has a cutoff of p value less than 0.00001. This was
 25 followed by annotation and literature searches to insure that the genes that passed the probability test have strong association with known inflammation genes. This process was reiterated so that the initial 41,419 genes were reduced to the final 11 inflammation-associated genes. Details of the expression patterns for the 11 novel inflammation genes are presented in Table 5.

Each of the 11 novel genes is coexpressed with at least one of the known genes with a p-
 30 value of less than 10E-05. The coexpression of the 11 novel genes with the known genes are shown in Table 5. The entries in Table 5 are the negative log of the p-value (- log p) for the coexpression of the two genes. The novel genes identified are listed in the table by their Incyte clone numbers, and the known genes their abbreviated names as shown in Ex.5. V

WO 00/29574

PCT/US99/26234

Table 5. Coexpression analysis of novel genes with known inflammation genes (-log p).

	CD16	L-selectin	SLAP	IP-30	p67-phox	AAT	p47-phox	C1qA	p40-phox	FLAP	HCK	1221361	3055142	402234	3507924	1335016	3054032	569989	2349263	2471716	2726173	706377	
5	CD16	4																					
	L-selectin	5	6																				
	SLAP	5	6																				
	IP-30	5	3	4																			
	P67-phox	3	5	7	1																		
	AAT	6	2	4	1	5																	
10	P47-phox	5	6	5	1	7	3	3															
	C1qA	7	2	6	1	3	7	6															
	P40-phox	5	5	3	4	7	4	6	1														
	FLAP	7	5	3	1	5	3	5	7	3													
	HCK	9	8	4	5	8	5	7	3	6	3												
	1221361	3	5	7	4	3	1	4	2	4	2	4											
15	3055142	6	5	9	4	5	3	8	1	4	3	8	4										
	402234	3	6	4	4	2	0	6	1	3	4	2	3	1									
	3507924	2	5	3	4	4	8	4	1	7	3	2	3	2	4								
	1335016	7	3	3	8	2	2	5	6	3	4	3	4	5	2	2							
20	3054032	5	11	9	5	4	4	7	7	7	8	7	5	7	5	4	4						
	569989	4	4	4	6	9	1	8	2	4	5	4	4	4	7	2	2	5					
	2349263	8	1	2	6	4	4	3	8	5	5	6	3	3	1	3	9	5	4				
	2471716	3	5	3	7	5	1	7	4	5	2	4	3	5	4	5	4	8	3	7			
	2726173	3	8	5	4	6	2	7	0	2	3	5	4	3	4	4	2	3	3	2	5		
25	706377	4	5	5	8	6	5	10	3	6	3	3	4	4	9	4	5	5	7	3	6	3	

V Novel Genes Associated with Inflammation

Eleven novel genes were identified from the data shown in Table 5 to be associated with inflammation.

Nucleic acids comprising the consensus sequences of SEQ ID NOs:1-11 of the present invention were first identified from Incyte Clones 402234, 569989, 706377, 1221361, 1335016, 2349263, 2471716, 2726173, 3054032, 3055142, and 3507924, respectively, and assembled according to Example III. BLAST and other motif searches were performed for SEQ ID NOs:1-11 according to Example VII. The sequences of SEQ ID NOs:1-11 were translated and sequence identity was sought with known sequences. Amino acids comprising the consensus sequences of SEQ ID NOs:12, 13, 14, 15, 16, and 17 of the present invention were encoded by the nucleic acids of SEQ ID NOs:1, 2, 6, 7, 8, and 11, respectively. SEQ ID NOs:12-17 were also analyzed using BLAST and other motif search tools as disclosed in Example VII.

SEQ ID NO:3 is 1229 nucleic acids in length and has about 99% match from about nucleic acid 250 to about nucleic acid 1216 with a human basement membrane-induced gene identified in

WO 00/29574

PCT/US99/26234

a human endometrial adenocarcinoma cell line (g3132521). SEQ ID NO:4 is about 1261 nucleic acids in length and has about 34% sequence identity from nucleic acid 23 to nucleic acid 994 with a predicted sequence similar to a RNA recognition motif (g2645068). SEQ ID NO:5 is 1340 nucleic acids in length and has about 60% sequence identity from about nucleic acid 21 to about 5 nucleic acid 925 with a human prostaglandin transporter hPGT mRNA (g3006201). The amino acid sequence encoded by SEQ ID NO:5 also exhibits several potential transmembrane domains as identified by HMM analysis. SEQ ID NO:9 is 2309 nucleic acids in length and shows similarity from about nucleic acid 104 to about nucleic acid 785 with a human polycystic kidney disease-associated protein gene (g790818).

10 SEQ ID NO:12 is 127 amino acid residues in length and shows about 50% sequence identity from about residue 37 to about residue 106 with a tobacco LIM-domain-containing protein (g1841464). The LIM domain is a cysteine-rich, zinc-binding motif of about 60 amino-acid residues that plays a potential role in DNA binding and regulation (Perez-Alvarado et al. (1994) Nat Struct Biol 1: 388-398). PFAM analysis shows that residues 40 to 97 of SEQ ID NO:12 15 encompasses the LIM domain. SEQ ID NO:13 is 93 amino acids in length and has a potential signal peptide sequence encompassing residues 1-18. SEQ ID NO:13 also exhibits a potential transmembrane domain encompassing residues 47 to 69. SEQ ID NO:14 is 225 amino acids in length and has about 32% sequence identity from about residue 5 to about residue 135 with a mouse high affinity IgE receptor beta subunit (g309225). SEQ ID NO:15 is 547 amino acid 20 residues in length and has about 35% sequence identity from about residue 413 to about 546 with a rat beta-chimaerin, a GTPase-activating protein expressed exclusively in the testis at the onset of sexual maturation (g203117). PFAM analysis shows that SEQ ID NO:15 has sequence homology from about residue 353 to about residue 523 with the GTPase-activator protein for Rho-like GTPases. SEQ ID NO:16 is 265 amino acids in length and shows about 93% sequence identity 25 from about residue 39 to about residues 265 with Maxp1, a rat protein which interacts with Ms4, a guanine nucleotide exchange factor (g2459833), and about 91% sequence identity from about residue 38 to about residue 265 with Nore1, a mouse putative Ras effector that plays an essential role in transmitting growth and differentiation signals received from Ras proteins(g2997698). This is confirmed by PFAM analysis that shows that SEQ ID NO:16 from about residue 119 to 30 about residue 211 matches a Ras association domain which interacts directly with the Ras proteins. SEQ ID NO:17 is 394 amino acids in length and exhibits a potential signal peptide sequence encompassing residues 1 to 19 and a potential transmembrane domain encompassing residues 273 to 295.

VI Homology Searching for Inflammation Genes and the Proteins

WO 00/29574

PCT/US99/26234

Polynucleotide sequences, SEQ ID NOs:1-11, and polypeptide sequences, SEQ ID NOs:12-17, were queried against databases derived from sources such as GenBank and SwissProt. These databases, which contain previously identified and annotated sequences, were searched for regions of similarity using BLAST (Altschul, *supra*) and Smith-Waterman alignment (Smith, *supra*). BLAST searched for matches and reported only those that satisfied the probability thresholds of 10^{-29} or less for nucleotide sequences and 10^{-8} or less for polypeptide sequences.

The polypeptide sequences were also analyzed for known motif patterns using MOTIFS, SPSCAN, BLIMPS, and Hidden Markov Model (HMM)-based protocols. MOTIFS (Genetics Computer Group, Madison WI) searches polypeptide sequences for patterns that match those defined in the Prosite Dictionary of Protein Sites and Patterns (Bairoch et al. *supra*), and displays the patterns found and their corresponding literature abstracts. SPSCAN (Genetics Computer Group) searches for potential signal peptide sequences using a weighted matrix method (Nielsen et al. (1997) *Prot Eng* 10:1-6). Hits with a score of 5 or greater were considered. BLIMPS uses a weighted matrix analysis algorithm to search for sequence similarity between the polypeptide sequences and those contained in BLOCKS, a database consisting of short amino acid segments, or blocks, of 3-60 amino acids in length, compiled from the PROSITE database (Henikoff and Henikoff, *supra*; Bairoch et al. *supra*), and those in PRINTS, a protein fingerprint database based on non-redundant sequences obtained from sources such as SwissProt, GenBank, PIR, and NRL-3D (Attwood et al. (1997) *J Chem Inf Comput Sci* 37:417-424). For the purposes of the present invention, the BLIMPS searches reported matches with a cutoff score of 1000 or greater and a cutoff probability value of 1.0×10^{-3} . HMM-based protocols were based on a probabilistic approach and searched for consensus primary structures of gene families in the protein sequences (Eddy, *supra*; Sonnhammer et al. *supra*). More than 500 known protein families with cutoff scores ranging from 10 to 50 bits were selected for use in this invention.

25 VII Labeling and Use of Individual Hybridization Probes

Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ - 32 P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (NEN Life Science Products, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine resin column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (NEN Life Science Products).

The DNA from each digest is fractionated on a 0.7 percent agarose gel and transferred to

WO 00/29574

PCT/US99/26234

NYTRANPLUS membranes (Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT AR film (Eastman Kodak, Rochester NY) is exposed to 5 the blots for several hours, hybridization patterns are compared.

VIII Production of Specific Antibodies

Any of the SEQ ID NOs:12-17, or fragments thereof, substantially purified using polyacrylamide gel electrophoresis (Harrington (1990) Methods Enzymol 182:488-495) or other purification techniques, is used to immunize rabbits and to produce antibodies using standard 10 protocols.

Alternatively, the amino acid sequence is analyzed using LASERGENE software (DNASTAR, Madison WI) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in 15 hydrophilic regions are well described in the art. Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A Peptide synthesizer (PE Biosystems) using Fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester to increase immunogenicity. Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for 20 antipeptide activity by, for example, by binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

WO 00/29574

PCT/US99/26234

What is claimed is:

1. A substantially purified polynucleotide comprising a gene that is coexpressed with one or more known inflammation genes in a plurality of biological samples, wherein each known inflammation gene is selected from the group consisting of CD16, L-selectin, Src-like adapter protein, IP-30, superoxidase homoenzyme subunits, alpha-1-antitrypsin, Clq-A, 5-lipoxygenase activating protein, and SRC family tyrosine kinase.
2. The polynucleotide of claim 1, comprising a polynucleotide sequence selected from:
 - (a) a polynucleotide sequence selected from the group consisting of SEQ ID NOs:1- 11;
 - (b) a polynucleotide sequence selected from the group consisting of SEQ ID NOs:12-17;
 - 10 (c) a polynucleotide sequence having at least 70% identity to the polynucleotide sequence of (a) or (b);
 - (d) a polynucleotide sequence which is complementary to the polynucleotide sequence of (a), (b), or (c);
 - (e) a polynucleotide sequence comprising at least 18 sequential nucleotides of the 15 polynucleotide sequence of (a),(b), (c), or (d); and
 - (f) a polynucleotide which hybridizes under stringent conditions to the polynucleotide of (a),(b),(c), (d) or (e).
3. A substantially purified polypeptide comprising the gene product of a gene that is coexpressed with one or more known inflammation genes in a plurality of biological samples, 20 wherein each known inflammation gene is selected from the group consisting of CD16, L-selectin, Src-like adapter protein, IP-30, superoxidase homoenzyme subunits, alpha-1-antitrypsin, Clq-A, 5-lipoxygenase activating protein, and SRC family tyrosine kinase.
4. The polypeptide of claim 3, comprising a polypeptide sequence selected from the group consisting of:
 - (a) the polypeptide sequence selected from the group consisting of SEQ ID NOs:12-17;
 - (b) a polypeptide sequence having at least 85% identity to the polypeptide sequence of 25 (a); and
 - (c) a polypeptide sequence comprising at least 6 sequential amino acids of the polypeptide sequence of (a) or (b).
- 30 5. An expression vector comprising the polynucleotide of claim 2.
6. A host cell comprising the expression vector of claim 5.
7. A pharmaceutical composition comprising the polynucleotide of claim 2 or the polypeptide of claim 3 in conjunction with a suitable pharmaceutical carrier.

WO 00/29574

PCT/US99/26234

8. An antibody which specifically binds to the polypeptide of claim 4.

9. A method for diagnosing a disease or condition associated with the altered expression of a gene that is coexpressed with one or more known inflammation genes, wherein each known inflammation gene is selected from the group consisting of CD16, L-selectin, Src-like adapter protein, IP-30, superoxidase homoenzyme subunits, alpha-1-antitrypsin, Clq-A, 5-lipoxygenase activating protein, and SRC family tyrosine kinase, the method comprising the steps of:

- (a) providing a sample comprising one of more of said coexpressed genes;
- (b) hybridizing the polynucleotide of claim 2 to said coexpressed genes under conditions effective to form one or more hybridization complexes;
- (c) detecting the hybridization complexes; and
- (d) comparing the levels of the hybridization complexes with the level of hybridization complexes in a non-diseased sample, wherein the altered level of hybridization complexes compared with the level of hybridization complexes of a nondiseased sample correlates with the presence of the disease or condition.

10. A method for treating or preventing a disease associated with the altered expression of a gene that is coexpressed with one or more known inflammation genes in a subject in need, wherein each known inflammation gene is selected from the group consisting of CD16, L-selectin, Src-like adapter protein, IP-30, superoxidase homoenzyme subunits, alpha-1-antitrypsin, Clq-A, 5-lipoxygenase activating protein, and SRC family tyrosine kinase, the method comprising the step of administering to said subject in need the pharmaceutical composition of claim 7 in an amount effective for treating or preventing said disease.

11. A method for treating or preventing a disease associated with the altered expression of a gene that is coexpressed with one or more known inflammation genes in a subject in need, wherein each known inflammation gene is selected from the group consisting of CD16, L-selectin, Src-like adapter protein, IP-30, superoxidase homoenzyme subunits, alpha-1-antitrypsin, Clq-A, 5-lipoxygenase activating protein, and SRC family tyrosine kinase, the method comprising the step of administering to said subject in need the antibody of claim 8 in an amount effective for treating or preventing said disease.

12. A method for treating or preventing a disease associated with the altered expression of a gene that is coexpressed with one or more known inflammation genes in a subject in need, wherein each known inflammation gene is selected from the group consisting of CD16, L-selectin, Src-like adapter protein, IP-30, superoxidase homoenzyme subunits, alpha-1-antitrypsin, Clq-A, 5-lipoxygenase activating protein, and SRC family tyrosine kinase, the method comprising the step of administering to said subject in need the polynucleotide sequence of claim 2 in an amount

2009年 6月12日 16時57分

S. YAMAMOTO OSAKA

NO. 8290 P. 48/71

WO 00/29574

PCT/US99/26234

effective for treating or preventing said disease.

WO 00/29574

PCT/US99/26234

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
 WALKER, Michael G.
 VOLKMUTH, Wayne
 KLINGLER, Tod M.

<120> INFLAMMATION-ASSOCIATED GENES

<130> PB-0006 PCT

<140> To Be Assigned
 <141> Herewith

<150> 09/195,292
 <151> 1998-11-18

<160> 17

<170> PERL Program

<210> 1
 <211> 1298
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No.: 402234CB1

<400> 1
 gggacatgac ttagggaaag tcccaaccgg aatcccccca gcccctgcct gtcaacaccc 60
 cccaccctgc aggctggggc cgggctggcg gggccctccc gactgacttc cccttgcaga 120
 acccaggcggt tgccgcttct ccacccggagg ctteccaccc tcacgagcca tggccaggc 180
 tgcaggagcc gcccaggcca cccctctca tgacgcctaa ggcggcggca gcagcacgg 240
 gcagcgctcc aagtccctca gcctgcgggc ccaggtgaag gagacctgcg cgcctgcca 300
 gaagaccgtg taccatgg agcggctggt ggccgacaag ctcattttcc acaaactttt 360
 ctctgctgc aagcaactgtc acaccaagct cagcctggc agctacgcgg cgctgcacgg 420
 ggagttctac tgcaaacccc acttccagca gctgttaag agcaaaggca actacgacga 480
 ggggtttggc cgcaagcagc acaaggagct ctgggcccac aaggagggtgg accccggcac 540
 caagacggcc tgaggccctct gtaaccttcc accccctctg cggaaaggcct ggagccggca 600
 gggggaaagg gggaaaggagg tcgagctggg ctgcgtggg ggcagggtgg gaagggatg 660
 aggcttgctc aggctgtggg gaccaggggca gggctctgtc ccaggactcc ttctttttc 720
 ctctcccgc agccggtagg gttttggaaa ccaggattgg ggtctgcucca ccaccctgtc 780
 tcctgcttcg ttcagccccc etccccaccc tcccccaggaa cccctggga ggcccccaag 840
 cccagctccc ctatcttagt gccttttcc cagcaaggag tcaagcatgcc cccctcagg 900
 tcccaagctc cctcaactgccc accggagact gtgtggccccc catgttcc catctaccc 960
 tacccttaac ctgtttctga gccacggaga cagggaggaa ggagcgccgac agtgcaccc 1020
 gttgggcata ataaatggcc ctgcagccca tgggggagga gatggggaaag tggagccacc 1080
 ctgcctctgc agggcaaggc agggcctgccc ccagtgggc ttgggaccat ctcgaaccac 1140
 cagcgtggag aaggcagaagc aaaagcactc gccaggctgc agcctcaggc actggcagg 1200
 gctgggtgcgg ccccaactccc ctccccccgtt cccattttgtt cccatccctgt tgtgaccaac 1260
 ccogttttaa acatgttca atagatccaa aaaaaaaa 1298

<210> 2
 <211> 532
 <212> DNA
 <213> Homo sapiens

WO 00/29574

PCT/US99/26234

<220>
<221> misc_feature
<223> Incyte ID No.: 569989CB1

<400> 2
ccccacgcgtc cgccttgaca ccagcagggt gacatccgct attgtactt ctctgctccc 60
ccacagtcc tctggacttc tctggaccac agtccctctgc cagacccctg ccagacccc 120
gtccaccatg atccatctgg gtcacatccct cttccatgttt ttgctcccaag tggctgcagc 180
tcagacgact ccaggagaga gatcatcaact coctgcctt taccctggca cttcaggctc 240
ttgttccgga tgtgggtccc tcctctgccc gtccttggca ggcctcggtt ctgtatgtc 300
ggtggcatcg ctgctcatcg tggggcggtt gttccatgtgc gcaegccac gccgagccc 360
cgcccaagaa gatggcaaaag tctacatcaa catgcaggc aggggctgac ctcctgcag 420
cttggacctt tgacttctga ccctctcatc ctggatgggt tggtggca caggaacccc 480
cgcccaact tttggattgt aataaaacaa ttgaaacacc aaaaaaaaaaa aa 532

<210> 3
<211> 1229
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 706377CB1

<400> 3
cagagagccg cggggaccat ggagccgggt cggctgcagg acttcgtgcg cgccctggac 60
ccccctccc tcccgccgt gctgccccgt tgcctgggg tctacttca gggccatccatc 120
tatgagatct ctgggaatga gtgtgcctc tccacggggg acctgtatca ggtcaccc 180
gtccgcctcc agaagggtgtt ctgtgagaac cccaaagacca gccagaccat ggagctcgcc 240
cccaacttcc aggtcttctc aagtcttagg attgcagcaac cacgctcgcc tgcccaaacc 300
caaggcgaag accttgcag agttcatcaa ggatggctcc agtacgtaca gcaagattcc 360
tgcccacagg aaggggccaca ggccgctaag ccccaaaaggc aggtatctaga tgatgtgaa 420
catgattatg aagaataact tggcaattt cagaaaaacca tctaagtgtt ggaggaacca 480
cgccctctaa ctgctgttcc tcaaggaaatc cgacaccagc caaccatcc aagcctctaa 540
aagacctcggt gcaagtctca cagaaactga gctgcagacg gggagtagct ttgtgaaac 600
tgattttagt gacactgcac cagttctt caggttcttag attcttgcata ctttagggcg 660
gttgttttgg acctaacatc tggcacgtga ctccctcagc ctcagaccc tggatgcag 720
agcagctggc agggtttcc tcaatccgtc aaccccaact gtcaccggg tggatgcaga 780
gggaaatccg aggccatcaa ccttgggtgac agcagcgcag tgccaatgtt gatcacactg 840
catgggat tttgttaacg tctgcaccc ccaactctcac ccccaagctc taagcccc 900
ggaggcctgg actgtctcc tcatctctgt agcaccaaagc ctgatagatc tttatatgg 960
aaacagggtt ttaaccatgtttaaca tggattaatg tggaaacttgc gttcaagaa 1020
cacaacctta ggacccttggg cccaaaaggc tgggttgaa atgaggagga gccaatttaa 1080
gaagaccctt atggagacatc gaggtgcag aaactgttag gttcatcag tgggttaag 1140
tcgtcaaagt tggatgtgac taaaccatgtt tatttccatcc taaaaccata gaataaaaaat 1200
gacacccgtt gttctctaaa aaaaaaaaaa 1229

<210> 4
<211> 1261
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1221361CB1

<400> 4
cggacgcgtg ggcggacgcg tggcggacgc gtggggcggac gctggggcgc gatggggctc 60
ttggaaacatt ggtgtgtca tcgtcattcg atgtgttgc atccacagga atctgggggt 120
gcacatatacc agggtaaatc cagttacact cggaccaggc actcaagaac agattcaatgt 180
catgcagag atggggaaatg gaaaggcataa ccgactttat gaagccatc tcctgagac 240

WO 00/29574

PCT/US99/26234

cttcggcga cctcagatag acccagctgt tgaaggattt attcgagaca aatatgagaa 300
 gaagaaaatac atggaccgaa gtcggacat caatgcctt aggaaagaaa aagatgacaa 360
 gtggaaaaga gggagcgaac cagttccaga aaaaaattg gaacctgttg ttttgagaa 420
 ggtaaaaatg ccacagaaaa aagaagaccc acagctaccc cggaaaagct ccccgaaatc 480
 cacagcgcct gtcatggatt ttttggcct tgatgcctt gtggcctgt ccattgcaas 540
 tagtaagacc agcaataccc tagagaagga tttagatctg ttggcctctg ttccatcccc 600
 ttcttcttcg gttccagaa agttgttagg ttccatgcc actgcaggga gtgcggctc 660
 tgtcctgaa aatctgaacc ttttccgga gccaggagc aaatcagaag aaataggcaa 720
 gaaacagctc tctaaagact ccattcttc actgtatgg tcccagacgc ctcaaatgcc 780
 tactcaagca atgttcatgg ctcccgcctca gatggcatat cccacagcct accccagctt 840
 cccgggggtt acacctccta acagcataat ggggagcatg atgcctccac cagtaggcat 900
 ggtgtctcg ccaggagctt ctggatggt tgcccccattg gccatgcctg caggttatat 960
 gggggcatg caggcatcaa tggatgggtgt gccaatggg atgatgacca cccagcaggg 1020
 tggctatcg cagggcatgg cagctatgcc ccagactgtg tatgggtcc agccagctca 1080
 gcagctgca tggAACCTTA ctcagatgac ccagcagatg gctggatgaa acttctatgg 1140
 agccaatggc atgatgaaact atggacagtc aatgagtggc ggaaatggac aggagcaaa 1200
 tcagactctc agtcctcaga tggtggataa aaaacaaaaac accttgtata aaaaaaaaaa 1260
 a 1261

<210> 5
<211> 1340
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1335016CB1

<400> 5
 ccctcgaaat tcggctcgag cagcaactcg cccctctacc tcgggatccct gtttgcagtg 60
 accatgatgg ggccaggcct ggcctttggg ctgggaccc tcatgcgtcg cctttatgtg 120
 gacattaacc agatgccaga aggtggatc agcctgacca taaaggaccc cccatgggtg 180
 ggtgcctgtt ggctgggtt cctcatcgct gcgggtgcag tgcccttggc tgccatcccc 240
 tacttcttc tcccaagga aatgccaag gaaaaacgtg agtttcgtt tcggcggaaag 300
 gtcttagcag tcacagactc acctgcccagg aagggaagg actctccctc taagcagagc 360
 cctggggagt ccacgaagaa gcaggatggc ctatccaga ttgcacccaa cctgactgtg 420
 atccagtta taaaagtctt ccccagggtg ctgctgcaga coctacgcca cccatcttc 480
 ctgctgggtgg tccctgtccca ggtatgcttgc tcateccatgg ctggggcat ggccacccctc 540
 ctgccccaaat tcctggagcg ccagtttcc atcacagctt cctacgccaa cctgtccatc 600
 gggtgcctct ccttccctc ggtcatcgat ggcacatgtgg tgggtggcgt cctggtaag 660
 cggtccacc tggggccctgt gggatgcggg qcctttggc tgctggggat gctgctgtgc 720
 ctcttcttca gcttgcgcgt ctcttttac ggctgcctca gccaccagat tgccggcattc 780
 acacaccaga ccagtgcaca ccctgggctg gagctgtetc caagctgcat ggaggcctgc 840
 tccctggccat tggacggctt taaccctgtc tgccacccca gcaactcgat ggaatacata 900
 acaccctgcc acgcaggatcg ctcaagctgg gtggccagg atgctctggc caacagccag 960
 agtccatccca cctcccccaccc tcattgcgtgg catcagcata taaacctgag gtcctccag 1020
 ggagagacct gggctgcact ggctgggtca gaagaacctg ttatgggtgc atagtccttc 1080
 agaagccagc caggcacac ctggggctga gagcccttcc agagaccccc aggccttggc 1140
 aggtggagca gtgaactctt gtggatatgg gaaccgattc aaatcccttct taggcctcta 1200
 actgactctg ttacctttagg caaattattt aactagtgc tcagttctt ggtctgtaaa 1260
 ataggggaga tattattaag tgctactac agagcaggaa tggctgtaaat aaatgttta 1320
 cctggatgaa aaaaaaaaaa 1340

<210> 6
<211> 2192
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2349263CB1

WO 00/29574

PCT/US99/26234

<400> 6

<210> 7

<211> 1992

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No.: 2471716CB1

<400> 7

aaaaaaactgt	gagagagaga	atttttaaaa	agcagctggg	gcctgagggtt	tctcccccag	60
tacccctgggt	cacccctcagcc	cagagctggc	ggcaggcccc	cagccctca	tgtcagagcc	120
ccctgtgtac	tgttaacctgg	tggacccctcg	ccgctgtct	cggtccccac	ccccagggcc	180
tgcatttttttt	ctgttgtcaga	ggctggatgc	ctgggagcag	cacctggacc	ccaactctgg	240
acgtctgttttc	tacataaaatt	caactgactgg	ctgcaagtcc	tggaagcccc	cgcccccgcag	300
tcgcagcggag	acgaaccctg	gctccatgg	ggggacacag	accctgaaga	ggaacaatga	360
tgttctgtcaa	cctcaggc	agggttcag	atctgacaca	gggaocccag	aaccgcttga	420
ccccacagggt	tcactcagcc	tcagccascg	cacccctgcag	cttgaccctc	cagccttgca	480
ggccccctcga	cctctggccgc	agctccctgg	cgaccccccatt	gagggtggaaa	agtccgggtct	540
gtctcaacatg	accaagattg	cccaaggggg	gcgcaagctc	aggaagaact	ggggccccgtc	600
ttgggtgttg	ttaacgggt	acagcttgg	gttctaccga	gagccaccgc	cgacagcgc	660
ctccctcaggc	tggggaccag	cgggttagccg	gccccgaaagt	agcgtggacc	tgcgcggggc	720
ggccctggcg	cacggccgco	acctgtccag	ccggccgcaac	gttctgcaca	tccgcacat	780

WO 00/29574

PCT/US99/26234

ccctggccac	gagttcctgc	tgcagtccga	ccacgagaca	gagctgcgag	cctggcaccg	840
cgcgtcgccc	actgtcatcg	agcggcttgg	tcggggagaac	ccccctggagc	tgcgctctgtc	900
gggcctcttgg	cccgccggagc	tgagogccgg	ggaggacgaa	gaagaggagt	cggagcttgt	960
gtccaaagccg	ctgtcgccgc	tcagcagccg	ccggagctcc	attcgggggc	ccgaaggcac	1020
cgagcagaac	cgcgtcgccg	acaaaactaaa	gccccctcatc	gcgaagagac	cgccttaca	1080
aaggcttcgag	gagcgggggtc	tgctccgaga	ccagggtgtc	ggctgccagt	tggaatcact	1140
ctgcccaggcg	gaaggagaca	cggtgcccag	ctttttgcgg	ctctgcattg	ctgctgtgga	1200
taaaagaggt	ctatgttgg	atggcattt	tcgggtgagc	gggaacttgg	cagtggtcca	1260
gaagcttcgc	tttctgggtg	acagagagcg	tcgggtcacc	tccgatggga	ggtatgtgtt	1320
ccoaagaacag	ccaggcacaag	aaggctcggt	agatttggac	agtactgagt	gggatgacat	1380
tcatgtggtc	accggagccc	tgaagcttt	tctccggggag	ctggcccccagc	ctctgggtcc	1440
accactgtcg	ctggcccccatt	tccgtgtgc	cettgcactc	tccgaatcag	agcagtgcct	1500
ctctcaaaaa	caagaattaa	taggtcaat	gccaaagccc	aaacctgaca	ctctacggta	1560
cctctggag	cattttatgc	gggtgatacg	acactcgat	aaaatcgca	tgacacccccca	1620
caacctggga	attgtgttt	gaccaaccct	gtttcgccca	gagcaggaga	cacatgcaccc	1680
agcagcccat	gctctcttacc	caggccagct	ggtccagctg	atgtccatca	acttcaccag	1740
cctcttcccc	tgatgcagg	aaggaaagaag	agaaaaacata	tttccggta	tctctgggtgg	1800
tgagaggctg	gtgttctgtt	ttgaggatat	cocttttaat	ctcccaaatg	actgtctcta	1860
tcttcatgag	tgtgacttga	ggtgttggga	ttgggtgaggg	agcttotctta	aagaggaaag	1920
tgagtggatt	aaaaacatcg	tctcttcttg	ttccctgtta	tcattcctcc	ccgaacataa	1980
taatacataa	gt					1992

<210> 8
<211> 3144
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2726173CB1

<400> 8	ccttgatgcg	ctggcggcct	cgccccggaa	ctccggggta	gatgaccgtg	gacagcagca	60
tgagcagtgg	gtactgcagc	ctggacgagg	aacttggaaa	ctgcttc	actgctaaga	120	
ctacctttt	cagaaatgcg	cagagcaaaac	atcttcaaa	aatgtctgt	aaacctgtgg	180	
aggagacaca	gcgcggcggc	acactgcagg	agatcaagca	gaagatcgac	agctacaaca	240	
cgcgagagaa	gaactgcctg	ggcatgaaac	tgagtgaaga	cggcacctac	acgggttca	300	
tcaaagtgc	tctgaaaactc	cgccggcttg	tgacgggtcc	tgctgggatc	cgccccca	360	
ccatctatga	tgccatcaag	gaggtgaacc	tggcggctac	cacggacaag	cgacatct	420	
tctacctgcc	cctagatgccc	atcaagcagc	tgcacatca	cagcaccacc	accgtcagt	480	
aggtcatcca	ggggctgtcc	aagaagttca	tgggtgtgg	caatccccag	aagttgcac	540	
tttttaagcg	gatacacaag	gacggacaag	tgctttcca	aaaactctcc	attgctgacc	600	
gcgcctctta	cctgcccctg	cttgctggc	ctgacacacgg	ggtcctcaac	tttgtgctaa	660	
aggagaatga	aactggagag	gttagagtgg	atgccttc	catccctgaa	cttcagaact	720	
tccttaacaat	cctggaaaaaa	gaggagcagg	acaaaatcca	acaagtgc	agaagtgatq	780	
acaaggtag	gcagaactg	gaggaggcct	taagagaatc	ccagggcaaa	cctgggtac	840	
cggctctgt	tccttcctcc	ctgggtcatt	cagattttat	tgtattatta	attattattt	900	
tgcacacagac	acttttttc	aggacatctc	tggcaggtgc	atttgtgcct	gcccagcagt	960	
tccagctgt	gcaaaaagtct	cttccatgg	caagtgtttt	cacgggggtt	cagctgtgcc	1020	
cgcccccaagg	ctgtggccca	ccacagattc	tgccaaaggat	cagaactcat	gtgaaacaaa	1080	
cagctgacgt	cctcttcg	totgeaagcc	tttcaccaac	caaatagtt	cctcttcgt	1140	
caccaaaactg	gaacotcaca	ccagccccca	aaggaaggaa	gaaaggtttt	agagctgtgt	1200	
gttctttctc	tggctttgtat	tettcttga	gttcttta	ttgcccacgt	caggaccatt	1260	
atttatgagt	gaaaagttgt	agcacattcc	tttgcaggt	ctgagacta	cccttgaag	1320	
cagggttaatg	ctcataaaaag	gactgttcc	gcccccccaa	ggtcctgtt	gttcacac	1380	
aagggaagtt	tataaagcta	ctggccccag	atgtcagg	taaggacac	caaagctgag	1440	
gctggctcag	agatctccag	agaagctgca	gcctgcctcg	gcccggc	tggccctgc	1500	
ccacattgca	catggaaaacc	caaaggcata	tatctgcgt	tgtgtggat	ttagtcat	1560	
tttgcac	aaactgttgc	tttttaagtt	acaaaatttga	atthaatgtt	gtcatcatcg	1620	
ccatgtgttt	ccccaaagggg	aagccagtca	ttgaccattt	aaaaagtctc	ctgctaagta	1680	
cgaaatcag	acagtaagag	aaagccaaaa	agcaatgcag	agaaagggtgt	ccaagctgtc	1740	
ttcagccttc	cccaagctaa	gagcagagga	ggccctggg	tacttgggtt	ccccatcg	1800	

WO 00/29574

PCT/US99/26234

ctccagcaact gctccatcc tcccactgcg actctggat ctccagggtgc tgcccaagga 1860
 gttgccttga ttacagagag gggagcctcc aattcgccca acttggatc ctttcgttt 1920
 tgaagcatgg gccagaccccg gcactgcgtc cggagagccg gtgggcctgg cctccccgtc 1980
 gacctcagtg ctttttgtt tttagagaga aataggatc gggcgagtt gcctgaagct 2040
 ctgtgtgtgg ctctctgtc caggaagtga acaatggccg cgggtgtggga gacaaggcca 2100
 ggagagcccg cgttcagtat gggttgaggg tcacagaccc cctccatc tgggtgcctg 2160
 agtttgact ccaatcgtg ataccagacc acatggacag ggaggatcaa attcctgact 2220
 tacatttgc cttggcttctt gtttaggtgt aatcctaaaaaa taataggatc aaaaaattcc 2280
 aacaagtgc caggactgc gagacactcc agtgcagagg gagaaggact tgtaatttcc 2340
 aaacgaggc tggtttcca acccagcctc tgagaaacca tttcttgc atccctgtcc 2400
 ttcccaagtc cctcttgggt cggtcaagc ccaagctgt tcgtgttagct tcagaagttc 2460
 cctctccgac ccaggctgag tcataactgc ccctgatccc agaaggaatg ctgacccttc 2520
 gtcgtatgaa ctgtcatacg tctccagagc ttcaaaaggca acacaagctc gcaactctaa 2580
 gatttttta aaccacaaaa accctggta gccatctcat gctcggccat atcacttccc 2640
 tccctttaga aactctctcc ctgctgtata tttaaggag cagggtggaga gtcattttcc 2700
 ttcgtctgc atgtctctaa cattaataga aggcatggct ctcgtgtccaa ccgctgtgaa 2760
 tgetgtgtgg aacccctcc tatggggatg gctattttat ttttggaaag gaaaaaaaaa 2820
 gtcatgtata tatacacata aaggcatata gctatataa aagagataag ggtgtttatg 2880
 aaatgagaaa attatggac aattcagact ttactaaagc acagtttagac ccaaggccct 2940
 tgctgaggc taaacctctg aaaaaagtat agtacgtact acccgttccc tcccgaggt 3000
 gggagtaact gctggtagtg cttctttgg ttgtgtgtc cagttgtgtaa gtgtttgtt 3060
 ccaggatatt ttctttttaa atgtcttct tatatgggtt ttaaaaaaaaa gtaataaaaag 3120
 cctgtgc aaatgaaaaa aaaa 3144

<210> 9
 <211> 2309
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No.: 3054032CB1

<400> 9
 aaggggccca ggaagatcaa gttgtgagg agaaatgggg aggaagttt cctgagtgcc 60
 tatgtgacc taagtccct tctggaccc aaaaaaaaaa tctggaaaggg ttcaggagg 120
 ctggaggggag aggcagcagg atgtggagg caggctctgg gacagggtgg ggaagagcag 180
 gcatgtctgg aagttgggg a gacaaagcag gctgagccgt gaggcaggct agacatcagg 240
 gaagaggcag agggaggcc aagacacaa gttggggctg gaaaggccag tgaggataga 300
 ggggaggctg ggggaaagcc aagacaaaaa gtcagatgtt gagaagggg tagggaaag 360
 acagaggccaa aggaagagaa gtcacaaagggt cagaagaagg ctgcacatgtt ggaggctaaa 420
 ggtgtggagg aaccaggagg agatgatgtt acagatgaga agggaaaaa aattgagaga 480
 gaagaggatg aacaaagaga ggaagcccag gtatagctg gaaggagccat agagcaagg 540
 gcccaggaaatc aatcaagttgc tgaggagaaa tggaaatgtt tacagaaaca agaggctgag 600
 ggagtctcagg aggtgaggaa caaaggacag agggagaagg ggtaccatgt agcaagaaaa 660
 gaccaaggaaatggtgaaga cagcagaagc ccagaagcag caactgaagg aggacccagg 720
 gaggtcagca aggaacgggaa ggtggggat ggagggctg agggagacca gagggtgg 780
 gggtactatt tagaaagagga caccctctt gaaagggtcg gtgtgtcgcc cctggagg 840
 gactgtgc aagaggccaa tcctcactt tctgagatgg aagaggtagc cccacagccaa 900
 cctcaggccataggatggaa gctgggggg cagccctgtc cagacggctg tctatcccc 960
 tggctcttgc gctggggatgg cgttggcatg cgttgcattt ccactctgtt tcaggccaa 1020
 caggccgtctgttggccccc aagccacatgt ttgccaagat gcccaggcc 1080
 atgtgttagca agattcatgtt ggcacactgoa aatccatgtcc cgaggccctgg ccggcttgat 1140
 gggactctgg gaaaaaggcc ttgggggtcc cgacttctc gatcctctt gggaaatggg 1200
 ggtgtcttgc tggatgttgc tggatgttgc ctggccggg accggccaaag gactgaggct 1260
 caaggatgttgc ggcgaacccaa gactgtact ggggtgggg attactgcct catccccaga 1320
 acctcccttgc ttagcatgtat cttccctggc cccttagctg cctggaggctc 1380
 ccacatgttgc gtcagaagg gttggatcc cggactgtcc ttagtgc ccccaagagaa 1440
 ccccaaggccatgttgc cttggatccat cagcggatcatatgcatt tgaaacacag 1500
 gctaaccctg gggaaagggtga aggactgtga ttaggaccac agggctggcc aaaggggacc 1560
 agcaagttgttgc tttgtatctc caggatgttgc gactgttgc tccctctgc gcatgaggc 1620
 ctgtgtgtcc caactctata ggtttggcc ctccaggcttcc tctctttgac tggggaggc 1680

WO 00/29574

PCT/US99/26234

actgccttgg ttggtttacc tgaacttgtc tccgacacaa agcaatttc tcttaggaga 1740
 tccccaaagaa agtcaacaag atcttgtcc caggagatgg gtcattggcc aaagggaaaca 1800
 taaggtggc agaaaactta aaagagttt ttaaagtgaa gactggagaa attcctccct 1860
 tcctctgac tgcataatc tttcatgaa agccaaagggt agagacagg aggacaggc 1920
 caggtaggg cctccacac aaaaacactt ctagatgtc ccattctgt tatgttctt 1980
 gaccctaaga tacctctgt ccctttaaa tccagattaa gagaaacgtc caggaagagc 2040
 tcttgaagc cctcaatatt tggaggg actggactcc tctccagctc cccaccctct 2100
 gcctccagtc accatgtgca agagaggtcc tgcatacgtc tctctggct ctccttctc 2160
 ctggata acttgcctt attcaggaa agggaaatgg tgcactcag gccctggac 2220
 tgcttctcca gccaggctgg gcccacaggt cccactctg tgaaggtaa tgtctcagaa 2280
 taaaagctgt attttacaa aaaaaaaaaa 2309

<210> 10
 <211> 1666
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No.: 3055142CB1

<400> 10
 ctgcctcgaga actgaatggc cctgtgcaga gccatagtcc cactgtgggt cctgcaatga 60
 gcaggggctg ggagtagagg gtttctgggg cctcagggtt ctggggaaagc aacagctatc 120
 agagagagaa gggccagacc ccatacgctc ttagatctt ggcagtagaa ggagaaggat 180
 gggtaatttgc acctctgaa tccctgacca ttgcattgtt cttagatctt ttctagaagg 240
 aagatctgag gctctgggtc tcaggggat ggcttggcc ttctctctca accttggctq 300
 agcctacccc ttaccttgc aaagacttgc ggaccctgtt tgcattgggt tgcattccct 360
 cctctgtggg gctcagggtt ttggaaatgtt gatgaaat ttctactt caagaccacc 420
 tctccctgca aacaccacac acacatggca tgcattgttgc cacatgcga cacacacacg 480
 cacacaccc aataattttt ctcaggatttt ctgatgttcc agaaaaaaacag cactaacgtt 540
 ggacctgtct actctcagaa cccggcacag attctctttt gatctctttt tggaaatctg 600
 gatttttaga agacaggata gggtaaaatt tagtagcgc tcaattttttt ctaaatctact 660
 agaggaagtt aattaaccc aaggcttcat ttctccagca ctaaaatggg gtggagagtt 720
 ggggtggaaa taagacatcc ttaaaagggtt aaattgtctg caaagcacct agcccaagtgc 780
 cgagctccca gttaggtttc agttaaaggat agtgcctgac ttctgttca cttttttttt 840
 ctgtttggat tcactggat actcttattt ccgttggat gtctcttact ctttccctgt 900
 tcgttggatgg ggcagaaacc agactgaaga gggaaagagac attccagaggg aggattgcct 960
 tcgtcagggt aagggttggg ctgtcagggtt gcocttccct tcaccccccctt ctgttatcaga 1020
 ttggccctcc cactccatc tcactctgtc tgcataatctt tccatatccq caagtttact 1080
 ggcactcttc tggcacctgg gcaagatccc agaacagagg atggagtgc tggccctcaca 1140
 gagcttagtg cccgactctg gggaaatggg actggggcat gggaaatggg cagecttagga 1200
 taggacacga gagtctgaaa ttcaaaagcaa ccagcttgc gttttttttt gatgttggaaag 1260
 caaacatggg ctagagagat agggcagaag tcaagacgag gatctggact gatgttggaga 1320
 aagttagccac ggaagcatga actgtatctt gcaaaatgc cctttttccc gcctccataat 1380
 tcattatgcc caaaaggct tacgttgcattt tccagccat agtactctg acttggagaga 1440
 cgtggacaga gccagcttctt accttgcctt gccgtcttc ccctgttctta atgtctgctc 1500
 ttgtctcaag ctccagaaga gtggccggcc atgtatctt aatatgtttt tggtgtatgg 1560
 gcaggttgc ttattatgtt atcaacagat gtccaggaac taatgagttt aatataat 1620
 tattgtcaaa taaaacttgc tttgttccat aaaaaaaaaa aaaaaa 1666

<210> 11
 <211> 1751
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No.: 3507924CB1

WO 00/29574

PCT/US99/26234

<400> 11

ttttagaggtt cctgtttgca tctctgcAAC cacttcagaa ggcacgtgtt tggtttgctc 60
 tgaggcataac cttagagtgtc cgccggcgtc ttcaattgtga gcttggggac tgcaatgtg 120
 gggagatttc agtgcatgtc cttccctggg tgctttcat cttggatttg aaagggtgaga 180
 gcacgtgtt ttgcccaactg aaactcatcc tgctgccagt gttactgttat tattccctgg 240
 gcctgaatga cttaaatgtt tccccgcctg agctaacaatgtt ccatgtgggt gattcaatgtc 300
 tggatggatg tggatggatg agcacacaaag acaaataatgtt attcaagata gactggactc 360
 tggtcaccagg agagcacggcc aaggacacaaatgtt atgtgtata ctattactcc aatctcagtg 420
 tggcttattgg gcccgttccag aaccgcgtac acctgtatggg ggacatcttta tgcaatgtg 480
 gctctcttctt gctccaaatgtt gtcaagaggc ctgaccaggc aacctatata tggtaaaatcc 540
 gcctcaaaagg ggagaggccag gtgttcaaga aggccgtgtt actgcgtatgtc cttccagagg 600
 agcccaaaga gtcatgttc catgtgggtt gattgttca gatggatgtt gtttccaga 660
 gcacagaagt gaaacacgtg accaaggatatt ttcaggacgg cggccaaagg 720
 aggagattgtt atttcggttac taccacaaaac tcaggatgtc tggatgttccatc 780
 gggccactt ccagaatgtc gtgaacccgt tggggacat tttccgttca gacgggttcca 840
 tcattgttca aggatgttggg gatgttccatc gaggaaacta caccgttcaatgttccatc 900
 ggaaaccttggt gttcaaaaaa accattgttc tggatgttca cccggaaaggag ccttcaacac 960
 tggtgacccc ggcagccctg agccctgttgg tcttgggtgg taatcgttgc gtgttcaatc 1020
 tggaaattgtt ctgttccaca atccctgttgc tccctgttgc gatattgttgc gtgaagaaga 1080
 cctgtggaaa taatgttca gtgaatttca cagtcttggt gaagaacacg aagaagacta 1140
 atccagagat aaaagaaaaa cccttccattt ttgaaagatgtt gtaaggggag aaacacatcc 1200
 actccccat aattgttacgg gaggtgtatcg aggaagaaga accaagtgtt aaatcagagg 1260
 ccacccatcat gaccatgttccatc ccaggccgttggc tttctgttgc gtcagatccg aacaacttcc 1320
 ttgaaaaaaa gtcaggtggg ggaatgtccaa aaacacacca agcccttgc gaaatgttcca 1380
 gagttcccttc atcttccatc cgggtggagac tcttcttgc gttgttccatc ggccacttca 1440
 ccagtgtttt cagactcccg ctctccatc tggatgttccatc tcttcttgc tggtaataac 1500
 actgaagatg gagaatttgg agccctgttgc agagactgtt gacgttgc ggaacacaggcc 1560
 tggatgggggg aggggagcat ggtttggcc tcttggatgtt gacactgttgc ctggaaacca 1620
 ggcttggatgtt agtggccatca aacccttgcgtt tggatgttccatc ccttccatc gcaatgttccatc 1680
 tagtggatgtt gttacttggg gaaatgttccatc ataaaaacca accccaaatca tttccctgttgc 1740
 aaaaaaaaaa a 1751

<210> 12

<211> 127

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No.: 402234CD1

<400> 12

Met	Phe	Gln	Ala	Ala	Gly	Ala	Ala	Gln	Ala	Thr	Pro	Ser	His	Asp
1														
														15
Ala	Lys	Gly	Gly	Gly	Ser	Ser	Thr	Val	Gln	Arg	Ser	Lys	Ser	Phe
														30
Ser	Leu	Arg	Ala	Gln	Val	Lys	Glu	Thr	Cys	Ala	Ala	Cys	Gln	Lys
														45
Thr	Val	Tyr	Pro	Met	Glu	Arg	Leu	Val	Ala	Asp	Lys	Leu	Ile	Phe
														60
His	Asn	Ser	Cys	Phe	Cys	Cys	Lys	His	Cys	His	Thr	Lys	Leu	Ser
														75
Leu	Gly	Ser	Tyr	Ala	Ala	Leu	His	Gly	Glu	Phe	Tyr	Cys	Lys	Pro
														90
His	Phe	Gln	Gln	Leu	Phe	Lys	Ser	Lys	Gly	Asn	Tyr	Asp	Glu	Gly
														105
Phe	Gly	Arg	Lys	Gln	His	Lys	Glu	Leu	Trp	Ala	His	Lys	Glu	Val
														120
Asp	Pro	Gly	Thr	Lys	Thr	Ala								
														125

WO 00/29574

PCT/US99/26234

<210> 13
<211> 93
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 569989CD1

<400> 13
Met Ile His Leu Gly His Ile Leu Phe Leu Leu Leu Pro Val
1 5 10 15
Ala Ala Ala Gln Thr Thr Pro Gly Glu Arg Ser Ser Leu Pro Ala
20 25 30
Phe Tyr Pro Gly Thr Ser Gly Ser Cys Ser Gly Cys Gly Ser Leu
35 40 45
Ser Leu Pro Leu Leu Ala Gly Leu Val Ala Ala Asp Ala Val Ala
50 55 60
Ser Leu Leu Ile Val Gly Ala Val Phe Leu Cys Ala Arg Pro Arg
65 70 75
Arg Ser Pro Ala Gln Glu Asp Gly Lys Val Tyr Ile Asn Met Pro
80 85 90
Gly Arg Gly

<210> 14
<211> 225
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2349263CD1

<400> 14
Met Thr Ser Gln Pro Val Pro Asn Glu Thr Ile Ile Val Leu Pro
1 5 10 15
Ser Asn Val Ile Asn Phe Ser Gln Ala Glu Lys Pro Glu Pro Thr
20 25 30
Asn Gln Gly Gln Asp Ser Leu Lys Lys His Leu His Ala Glu Ile
35 40 45
Lys Val Ile Gly Thr Ile Gln Ile Leu Cys Gly Met Met Val Leu
50 55 60
Ser Leu Gly Ile Ile Leu Ala Ser Ala Ser Phe Ser Pro Asn Phe
65 70 75
Thr Gln Val Thr Ser Thr Leu Leu Asn Ser Ala Tyr Pro Phe Ile
80 85 90
Gly Pro Phe Phe Phe Ile Ile Ser Gly Ser Leu Ser Ile Ala Thr
95 100 105
Glu Lys Arg Leu Thr Lys Leu Leu Val His Ser Ser Leu Val Gly
110 115 120
Ser Ile Leu Ser Ala Leu Ser Ala Leu Val Gly Phe Ile Ile Leu
125 130 135
Ser Val Lys Gln Ala Thr Leu Asn Pro Ala Ser Leu Gln Cys Glu
140 145 150
Leu Asp Lys Asn Asn Ile Pro Thr Arg Ser Tyr Val Ser Tyr Phe
155 160 165
Tyr His Asp Ser Leu Tyr Thr Thr Asp Cys Tyr Thr Ala Lys Ala
170 175 180
Ser Leu Ala Gly Thr Leu Ser Leu Met Leu Ile Cys Thr Leu Leu

WO 00/29574

PCT/US99/26234

Glu	Phe	Cys	Leu	Ala	Val	Leu	Thr	Ala	Val	Leu	Arg	Trp	Lys	Gln
				185						190				195
					200				205					210
Ala	Tyr	Ser	Asp	Phe	Pro	Gly	Val	Ser	Val	Leu	Ala	Gly	Phe	Thr
									215					225
										220				

<210> 15
<211> 547
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2471716CD1

<400>	15													
Met	Ser	Glu	Pro	Pro	Val	Tyr	Cys	Asn	Leu	Val	Asp	Leu	Arg	Arg
1		5							10				15	
Cys	Pro	Arg	Ser	Pro	Pro	Pro	Gly	Pro	Ala	Cys	Pro	Leu	Leu	Gln
					20				25				30	
Arg	Leu	Asp	Ala	Trp	Glu	Gln	His	Leu	Asp	Pro	Asn	Ser	Gly	Arg
					35				40				45	
Cys	Phe	Tyr	Ile	Asn	Ser	Leu	Thr	Gly	Cys	Lys	Ser	Trp	Lys	Pro
					50				55				60	
Pro	Arg	Arg	Ser	Arg	Ser	Glu	Thr	Asn	Pro	Gly	Ser	Met	Glu	Gly
					65				70				75	
Thr	Gln	Thr	Leu	Lys	Arg	Asn	Asn	Asp	Val	Leu	Gln	Pro	Gln	Ala
					80				85				90	
Lys	Gly	Phe	Arg	Ser	Asp	Thr	Gly	Thr	Pro	Glu	Pro	Leu	Asp	Pro
					95				100				105	
Gln	Gly	Ser	Leu	Ser	Leu	Ser	Gln	Arg	Thr	Ser	Gln	Leu	Asp	Pro
					110				115				120	
Pro	Ala	Leu	Gln	Ala	Pro	Arg	Pro	Leu	Pro	Gln	Leu	Leu	Asp	Asp
					125				130				135	
Pro	His	Glu	Val	Glu	Lys	Ser	Gly	Leu	Leu	Asn	Met	Thr	Lys	Ile
					140				145				150	
Ala	Gln	Gly	Gly	Arg	Lys	Leu	Arg	Lys	Asn	Trp	Gly	Pro	Ser	Trp
					155				160				165	
Val	Val	Leu	Thr	Gly	Asn	Ser	Leu	Val	Phe	Tyr	Arg	Glu	Pro	Pro
					170				175				180	
Pro	Thr	Ala	Pro	Ser	Ser	Gly	Trp	Gly	Pro	Ala	Gly	Ser	Arg	Pro
					185				190				195	
Glu	Ser	Ser	Val	Asp	Leu	Arg	Gly	Ala	Ala	Leu	Ala	His	Gly	Arg
					200				205				210	
His	Leu	Ser	Ser	Arg	Arg	Asn	Val	Leu	His	Ile	Arg	Thr	Ile	Pro
					215				220				225	
Gly	His	Glu	Phe	Leu	Leu	Gln	Ser	Asp	His	Glu	Thr	Glu	Leu	Arg
					230				235				240	
Ala	Trp	His	Arg	Ala	Leu	Arg	Thr	Val	Ile	Glu	Arg	Leu	Asp	Arg
					245				250				255	
Glu	Asn	Pro	Leu	Glu	Leu	Arg	Leu	Ser	Gly	Ser	Gly	Pro	Ala	Glu
					260				265				270	
Leu	Ser	Ala	Gly	Glu	Asp	Glu	Glu	Glu	Ser	Glu	Leu	Val	Ser	
					275				280				285	
Lys	Pro	Leu	Leu	Arg	Leu	Ser	Ser	Arg	Arg	Ser	Ser	Ile	Arg	Gly
					290				295				300	
Pro	Glu	Gly	Thr	Glu	Gln	Asn	Arg	Val	Arg	Asn	Lys	Leu	Lys	Arg
					305				310				315	
Leu	Ile	Ala	Lys	Arg	Pro	Pro	Leu	Gln	Ser	Leu	Gln	Glu	Arg	Gly
					320				325				330	
Leu	Leu	Arg	Asp	Gln	Val	Phe	Gly	Cys	Gln	Leu	Glu	Ser	Leu	Cys
					335				340				345	

WO 00/29574

PCT/US99/26234

Gln	Arg	Glu	Gly	Asp	Thr	Val	Pro	Ser	Phe	Leu	Arg	Leu	Cys	Ile
350									355					360
Ala	Ala	Val	Asp	Lys	Arg	Gly	Leu	Asp	Val	Asp	Gly	Ile	Tyr	Arg
365									370					375
Val	Ser	Gly	Asn	Leu	Ala	Val	Val	Gln	Lys	Leu	Arg	Phe	Leu	Val
380									385					390
Asp	Arg	Glu	Arg	Ala	Val	Thr	Ser	Asp	Gly	Arg	Tyr	Val	Phe	Pro
395									400					405
Glu	Gln	Pro	Gly	Gln	Glu	Gly	Arg	Leu	Asp	Leu	Asp	Ser	Thr	Glu
410									415					420
Trp	Asp	Asp	Ile	His	Val	Val	Thr	Gly	Ala	Leu	Lys	Leu	Phe	Leu
425									430					435
Arg	Glu	Leu	Pro	Gln	Pro	Leu	Val	Pro	Pro	Leu	Leu	Leu	Pro	His
440									445					450
Phe	Arg	Ala	Ala	Leu	Ala	Leu	Ser	Glu	Ser	Glu	Gln	Cys	Leu	Ser
455									460					465
Gln	Ile	Gln	Glu	Leu	Ile	Gly	Ser	Met	Pro	Lys	Pro	Asn	His	Asp
470									475					480
Thr	Leu	Arg	Tyr	Leu	Leu	Glu	His	Leu	Cys	Arg	Val	Ile	Ala	His
485									490					495
Ser	Asp	Lys	Asn	Arg	Met	Thr	Pro	His	Asn	Leu	Gly	Ile	Val	Phe
500									505					510
Gly	Pro	Thr	Leu	Phe	Arg	Pro	Glu	Gln	Glu	Thr	Ser	Asp	Pro	Ala
515									520					525
Ala	His	Ala	Leu	Tyr	Pro	Gly	Gln	Leu	Val	Gln	Leu	Met	Leu	Thr
530									535					540
Asn	Phe	Thr	Ser	Leu	Phe	Phe								
				545										

<210> 16
<211> 265
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2726173CD1

<400> 16															
Met	Thr	Val	Asp	Ser	Ser	Met	Ser	Ser	Gly	Tyr	Cys	Ser	Leu	Asp	
1						5			10					15	
Glu	Glu	Leu	Glu	Asp	Cys	Phe	Phe	Thr	Ala	Lys	Thr	Thr	Phe	Phe	
						20			25					30	
Arg	Asn	Ala	Gln	Ser	Lys	His	Leu	Ser	Lys	Asn	Val	Cys	Lys	Pro	
						35			40					45	
Val	Glu	Glu	Thr	Gln	Arg	Pro	Pro	Thr	Leu	Gln	Glu	Ile	Lys	Gln	
						50			55					60	
Lys	Ile	Asp	Ser	Tyr	Asn	Thr	Arg	Glu	Lys	Asn	Cys	Leu	Gly	Met	
						65			70					75	
Lys	Leu	Ser	Glu	Asp	Gly	Thr	Tyr	Thr	Gly	Phe	Ile	Lys	Val	His	
						80			85					90	
Leu	Lys	Leu	Arg	Arg	Pro	Val	Thr	Val	Pro	Ala	Gly	Ile	Arg	Pro	
						95			100					105	
Gln	Ser	Ile	Tyr	Asp	Ala	Ile	Lys	Glu	Val	Asn	Leu	Ala	Ala	Thr	
						110			115					120	
Thr	Asp	Lys	Arg	Thr	Ser	Phe	Tyr	Leu	Pro	Leu	Asp	Ala	Ile	Lys	
						125			130					135	
Gln	Leu	His	Ile	Ser	Ser	Thr	Thr	Thr	Val	Ser	Glu	Val	Ile	Gln	
						140			145					150	
Gly	Leu	Leu	Lys	Lys	Phe	Met	Val	Val	Asp	Asn	Pro	Gln	Lys	Phe	
						155			160					165	
Ala	Leu	Phe	Lys	Arg	Ile	His	Lys	Asp	Gly	Gln	Val	Leu	Phe	Gln	

WO 00/29574

PCT/US99/26234

Lys	Leu	Ser	Ile	Ala	Asp	Arg	Pro	Leu	Tyr	Leu	Arg	Leu	Leu	Ala
170				185					175					180
Gly	Pro	Asp	Thr	Glu	Val	Leu	Asn	Phe	Val	Leu	Lys	Glu	Asn	Glu
				200					205					195
Thr	Gly	Glu	Val	Glu	Trp	Asp	Ala	Phe	Ser	Ile	Pro	Glu	Leu	Gln
				215					220					210
Asn	Phe	Leu	Thr	Ile	Leu	Glu	Lys	Glu	Gln	Asp	Lys	Ile	Gln	225
				230					235					240
Gln	Val	Gln	Lys	Lys	Tyr	Asp	Lys	Phe	Arg	Gln	Lys	Leu	Glu	Glu
				245					250					255
Ala	Leu	Arg	Glu	Ser	Gln	Gly	Lys	Pro	Gly					
				260					265					

<210> 17
<211> 394
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 3507924CD1

<400> 17

Met	Phe	Cys	Pro	Leu	Lys	Leu	Ile	Leu	Leu	Pro	Val	Leu	Leu	Asp
1				5					10					15
Tyr	Ser	Leu	Gly	Leu	Asn	Asp	Leu	Asn	Val	Ser	Pro	Pro	Glu	Leu
				20					25					30
Thr	Val	His	Val	Gly	Asp	Ser	Ala	Leu	Met	Gly	Cys	Val	Phe	Gln
				35					40					45
Ser	Thr	Glu	Asp	Lys	Cys	Ile	Phe	Lys	Ile	Asp	Trp	Thr	Leu	Ser
				50					55					60
Pro	Gly	Glu	His	Ala	Lys	Asp	Glu	Tyr	Val	Leu	Tyr	Tyr	Tyr	Ser
	65								70					75
Asn	Leu	Ser	Val	Pro	Ile	Gly	Arg	Phe	Gln	Asn	Arg	Val	His	Leu
	80								85					90
Met	Gly	Asp	Ile	Leu	Cys	Asn	Asp	Gly	Ser	Leu	Leu	Leu	Gln	Asp
	95								100					105
Val	Gln	Glu	Ala	Asp	Gln	Gly	Thr	Tyr	Ile	Cys	Glu	Ile	Arg	Leu
	110								115					120
Lys	Gly	Glu	Ser	Gln	Val	Phe	Lys	Lys	Ala	Val	Val	Leu	His	Val
	125								130					135
Leu	Pro	Glu	Glu	Pro	Lys	Glu	Leu	Met	Val	His	Val	Gly	Gly	Leu
	140								145					150
Ile	Gln	Met	Gly	Cys	Val	Phe	Gln	Ser	Thr	Glu	Val	Lys	His	Val
	155								160					165
Thr	Lys	Val	Glu	Trp	Ile	Phe	Ser	Gly	Arg	Arg	Ala	Lys	Glu	Glu
	170								175					180
Ile	Val	Phe	Arg	Tyr	Tyr	His	Lys	Leu	Arg	Met	Ser	Val	Glu	Tyr
	185								190					195
Ser	Gln	Ser	Trp	Gly	His	Phe	Gln	Asn	Arg	Val	Asn	Leu	Val	Gly
	200								205					210
Asp	Ile	Phe	Arg	Asn	Asp	Gly	Ser	Ile	Met	Leu	Gln	Gly	Val	Arg
	215								220					225
Glu	Ser	Asp	Gly	Gly	Asn	Tyr	Thr	Cys	Ser	Ile	His	Leu	Gly	Asn
	230								235					240
Leu	Val	Phe	Lys	Lys	Thr	Ile	Val	Leu	His	Val	Ser	Pro	Glu	Glu
	245								250					255
Pro	Arg	Thr	Leu	Val	Thr	Pro	Ala	Ala	Leu	Arg	Pro	Leu	Val	Leu
	260								265					270
Gly	Gly	Asn	Gln	Leu	Val	Ile	Ile	Val	Gly	Ile	Val	Cys	Ala	Thr

WO 00/29574

PCT/US99/26234

	275		280		285
Ile Leu Leu Leu	Pro Val Leu Ile Leu	Ile	Val Lys Lys Thr	Cys	
290	295		300		
Gly Asn Lys Ser	Ser Val Asn Ser Thr	Val	Leu Val Lys Asn	Thr	
305	310		315		
Lys Lys Thr Asn	Pro Glu Ile Lys Glu	Lys	Pro Cys His Phe	Glu	
320	325		330		
Arg Cys Glu Gly	Glu Lys His Ile Tyr	Ser	Pro Ile Ile Val	Arg	
335	340		345		
Glu Val Ile Glu	Glu Glu Pro Ser	Glu	Lys Ser Glu Ala	Thr	
350	355		360		
Tyr Met Thr Met	His Pro Val Trp Pro	Ser	Leu Arg Ser Asp	Arg	
365	370		375		
Asn Asn Ser Leu	Glu Lys Lys Ser Gly	Gly	Gly Met Pro Lys	Thr	
380	385		390		
Gln Gln Ala Phe					